PAUL POSTAL

Constituent Structure: A Study of Contemporary Models of Syntactic Description

INDIANA UNIVERSITY, BLOOMINGTON
MOUTON & CO., THE HAGUE, THE NETHERLANDS

JOHN G. McOKE
7340 - 24 TH AVE. N. E.
SEATTLE, WASH. 98115
23 Sept 69

CONSTITUENT STRUCTURE: A STUDY OF CONTEMPORARY MODELS OF SYNTACTIC DESCRIPTION

On loan to A216

RESEARCH CENTER FOR THE LANGUAGE SCIENCES

INDIANA UNIVERSITY

Thomas A. Sebeok *Chairman*

Andrew Vázsonyi *Associate Chairman, Publications*

Constituent Structure: A Study of Contemporary Models of Syntactic Description

by

PAUL POSTAL

Published by
INDIANA UNIVERSITY, BLOOMINGTON
MOUTON & CO., THE HAGUE, THE NETHERLANDS

Publication Thirty

(Third Edition)

This book was originally published as
Part II of the International Journal of American Linguistics
Volume 30, No. 1 (1964)

Library of Congress Catalog Card Number: 63-64432

All orders from the United States of America and from Canada should be addressed to Publications, Research Center for the Language Sciences, Patton House, Indiana University, Bloomington, Indiana 47401. Orders from all other countries should be sent to Mouton & Co., Publishers, The Hague, The Netherlands.

Printed in the United States of America

PREFACE

Within the last half dozen years an increasing amount of work on the grammatical description of natural languages has been done within the framework of 'generative grammar'. This framework is methodological. It represents a proposal about the way linguistic research should proceed, the aims this should take, etc. This framework is empirically neutral and excludes no possible claim about the nature of language. Within this framework there has developed, however, a substantive theory of grammar based largely on the notion of 'transformational rule'. In much of the work on transformational grammar there is the claim, implicit or explicit, that another theory of language, the so-called theory of phrase structure, which was developed in precise formal terms only within the framework of generative grammar, correctly represents the underlying conception of grammatical description prevelant in modern linguistics generally, especially in modern American linguistics. A good deal of the research relating to transformational grammars has been concerned to show the inadequacy of the theory of phrase structure, especially its inadequacy vis a vis the theory of transformational grammar.

Yet it must be noted that although there is, to my knowledge, no published work which explicitly disputes the counterevidence to the theory of phrase structure which has been brought forward in the course of work on transformational grammars, nearly every study of either syntactic theory or of the actual morphosyntactic properties of individual languages which both lies within the tradition of modern linguistics and which has been published during the six year period since the development of transformational grammars, remains wholly within the framework which most followers of transformational grammar believe to be characterized by the formal theory of phrase structure. This extensive body of linguistic work thus represents a massive though implicit denial of the claims emenating from research on transformational grammar and suggests either or both of the following propositions:

(1) The majority of linguists are unconvinced of the inadequacy of the theory of phrase structure as a model of human language.

(2) The majority of linguists are unconvinced that the theory of phrase structure correctly represents the morphosyntactic ideas with which most modern linguists work.

From the point of view of those who believe that the claims implicit in work on transformational grammars are correct, the present situation in linguistics suggests that it is necessary to try and show more clearly what has been claimed in the past by advocates of this theory with respect to the prevailing conception of the form of grammar, namely, that the theory of phrase structure does represent modern ideas of morphosyntactic description and that this theory is inadequate. This is the motivation which prompts the present monograph and its three basic aims. The first of these is to describe in some detail the theory of phrase structure. The second and by far the most important is to show that modern American syntactic ideas are correctly represented, formalized, or explicated by this theory. Naturally this cannot be done by dealing with every explicit statement about the nature of grammar or every instance of actual language description. I have therefore picked for consideration what I take to be the most representative, influential, and presently most important and widely accepted views. I do not think these choices can be regarded as controversial. The third aim is to consider the correctness of the theory of phrase structure. Besides briefly reviewing evidence of more well known varieties, we shall have some new and very important evidence of a hitherto unavailable type to present which bears on the truth of this theory.

If the present study is successful, I hope to have simplified the choices available to the student of grammar who is concerned with determining the relative truth values of the various apparently distinct views of grammar which have been proposed within modern linguistics. If the argument of this work is sound, there are really only two: transformational syntax and the theory of phrase structure, and the evidence relevant to the choice between them is clear and decisive.

The research embodied in this monograph was supported by the U. S. Army, the Air Force Office of Scientific Research, and the Office of Naval Research. I am indebted to Noam Chomsky for extensive criticism of earlier versions of this work and for many particular insights into its subject matter not specifically noted below. The wider and more general debt I owe him is evident throughout. I wish to thank as well Morris Halle, Jerrold Katz, and G. H. Matthews, all of whom read earlier versions of this work and who are responsible for many improvements in content and style.

TABLE OF CONTENTS

Chapter 1: Background 1
I. Introduction 1
II. Grammatical Descriptions and Grammatical Theory 3
A. Requirements for Particular Grammatical Descriptions, ... 3
B. Requirements for General Linguistic Theory. 3
Chapter 2: Phrase Markers 6
Chapter 3: Rewriting Systems 9
Chapter 4: Some Proposed Models of Grammatical Description . . 18
I. Remarks 18
II. Bloch's Japanese Syntax. 18
III. Wells' Immediate Constituent Approach 21
IV. Harris' Morpheme Class Substitution System. 25
V. Hockett's Item and Arrangement System 28
VI. Lamb's Stratificational Syntax. 30
VII. Tagmemics 33
VIII. Agreement, Variables, and PSG Rules 43
IX. Tagmemics: Conclusions 50
X. Hockett's Constructional Grammar 51
XI. String Analysis. 61
Chapter 5: Discontinuities, Rewrite Rules, and P-markers . . . 67
Chapter 6: Other Systems. 71
Chapter 7: The Inadequacies of PSG 72
Chapter 8: Conclusion 78
I. Substantive Conclusions 78
II. Methodological Conclusions 79
Notes 81
Appendix: Halliday's Categories of the Theory of Grammar . . 97
Bibliography 118

CHAPTER 1: BACKGROUND

I. Introduction

Superficially, there appear to have developed since the Second World War a wide variety of distinct models for the syntactic description of natural languages. Taking only the writings of American linguists,[1] one finds among the major approaches such apparently different conceptions as Harris' 'morpheme to utterance procedures'; the traditional immediate constituent approach perhaps most closely associated with Bloch, Wells, and Hockett; the tagmemic model of Pike, Longacre, and their associates of the Summer Institute of Linguistics; Lamb's 'stratificational grammar'; Hockett's 'constructional grammar'; Harris' 'string analysis'; and finally the quite distinct transformational conceptions of Harris on the one hand and Chomsky on the other.[2] The purpose of the present study is to argue that despite a multitude of minor differences, both of substance and terminology, there is an essential equivalence among those syntactic conceptions prevelant in the United States which do not utilize transformational devices.[3] This will be argued by showing that these conceptions of grammar are, with certain provisos, versions of the theory of phrase structure grammars in the sense of Chomsky.[4] This demonstration, if successful, will bear heavily on the question of the relative adequacy of TG with respect to available nontransformational alternatives.

Four major questions will be discussed:

1. What is a PSG?
2. What is the relation between current nontransformational syntactic ideas and the theory of PSG? To what extent are the formulations of these ideas versions of PSG?
3. What is the status of the theory of PSG with respect to empirical adequacy? That is, what is known about the truth of this theory?

4. What can be said about the empirical adequacy of any features of prevelant nontransformational syntactic views which lie outside the domain of PSG?

Interest in these questions leads to a brief preliminary discussion of the requirements which individual grammars and general linguistic theory must meet, and to scattered comparisons between TG and the major nontransformational approaches which are considered.

American descriptive grammar is firmly rooted in what Chomsky[5] has called the 'taxonomic' view of linguistics. That is, the major motivation which has influenced the form of modern descriptive grammar appears to be the belief that some sort of taxonomic methodology, based on segmentation and classification of utterances and their parts, can be formulated which will derive correct grammatical descriptions from the data. The restriction that grammatical description may contain only devices which can be discovered by such methods imposes essential restrictions on the form of grammar. The result is a conception of sentence structure succinctly characterized by Lees:[6]

> '. . . the linguist, . . . has assumed that the sentences of a language may each be analyzed into a linearly concatenated sequence of immediate constituents, and that this bracketing or parsing operation may be performed at various levels of generality to yield a hierarchical branching-diagram, such that any unit at any level is just a certain continuous string within some sentence or else a class of such strings drawn from different but grammatically equivalent sentences.'

and a conception of grammar described by Chomsky as 'taxonomic':[7]

> 'Each rule is of the form: element A has the member (variant, realization) X in the context Z__W. Let us call such a rule a rewriting rule. The syntactic component consists of an unordered set of rewriting rules, each of which states the membership of some phrase category or formative category in some context. The structural description that it provides can be regarded as a labelled bracketing of the string of formatives.'

That is, the fundamental idea in taxonomic grammar is that a linguistic description on the syntactic level is a kind of pyramid of constituents with the apex being the highest order constituent, and each successively lower level giving the set of possible sequences of the immediately higher constituent. A full grammar is then the maximally general or complete pyramid showing all expansions for

all constituents; the structural description of each sentence is a special case of this with a particular specified selection of the different expansion possibilities at each level. PSG are simply the maximally general formulation of the ideas involved in this 'pyramid' type of grammatical model. Before describing the theory of PSG more precisely, it will be helpful to describe briefly the context in which theories of grammar must be considered.

II. Grammatical Descriptions and Grammatical Theory

A. Requirements for Particular Grammatical Descriptions. I shall assume in the following that the minimal requirements for a grammatical description of a natural language are that it describe precisely what the sentences of a language are and that it say precisely what structure each has. In other words, a grammar must be an explicit formal device which enumerates all and only the well-formed strings and which automatically assigns to each sentence a correct structural description (henceforth SD) showing what elements the sentence contains, their relations to each other, the relations of the sentence to other sentences, etc.[8] The SD of a sentence must provide an account of all grammatical information in principle available to the native speaker. Thus a grammar of English should explicitly inform us that John likes meat is a sentence, that likes not runs off is not; that John and the body are, on some level (say Noun Phrase), expressions of the same type as opposed to likely, or eats fast; that the killing of the tigers is ambiguous, with one interpretation in which tigers is the 'subject' of killing as in tigers kill, but with another in which tigers is the 'object' of killing as in they kill tigers; that I will not go home is the 'negation' of I will go home while will I go home is the 'yes-no question' version of the former; etc.

B. Requirements for General Linguistic Theory. Grammatical descriptions of individual languages must provide an explicit enumeration of pairs of grammatical utterances (sentences, well-formed strings) and their correct associated SD. A theory about the nature of language, a 'linguistic theory', on the other hand, must provide a general account of the universal features of all grammars. It must then contain at least the following:

1. a precise characterization of the possible types of grammatical rule and their possible interrelations.

2. a characterization of the kinds of SD.

3. a mechanical procedure (algorithm) for associating a unique SD with each enumerated sentence.

4. an evaluation procedure or metric of simplicity for grammars to choose the best grammar out of all those compatible with the data.

The conditions insisted on for particular grammatical descriptions and for general linguistic theory are the weakest possible if linguistics is to provide an account of the available data. Consider first grammars of particular languages. Failure to generate all the sentences means the grammar does not describe the full language. Failure to generate only sentences means the grammar fails to match the native speaker's fundamental ability to discriminate sentence from nonsentence. Failure to automatically and uniquely assign SD to enumerated sentences means the grammar does not make clear what claims it makes about each sentence and thus denies the grammar much of its testability. Assignment of incorrect SD means that false claims are made about part of the subject matter. Failure to make the rules of the grammar explicit and precise makes it impossible to determine what sentences the grammar does or does not enumerate and just what structures it claims that these sentences have. To just the extent of its imprecision, a grammar cannot be checked against empirical linguistic data.

On the other hand, failure on the part of linguistic theory to specify precisely the possible types of linguistic rule and the possible types of SD, automatically means the theory does not characterize the notion 'grammar of a language'. Similarly, if no algorithm of structure assignment is given for associating an SD with each sentence, grammars drawn from the theory cannot provide an explanation of why sentences have the kinds of structures they do. Put differently, the algorithm of structure assignment is required to show why a person who internalizes just the types of grammatical rules provided by the theory would recognize that sentences have the appropriate grammatical analyses. Only a general theory which provides a mechanism for associating grammatical analyses with the output of grammatical rules can make any specific nonarbitrary claims about the structure of sentences. I have belabored this point because it is here that there are the greatest gaps in modern descriptive grammar. In general, linguistic theory must be precisely formulated so that given some particular grammatical

description which is either confirmed or disconfirmed, we may tell whether this is in fact confirmatory or not for the general theory. Unless this requirement is met, investigations into the structure of particular languages cannot be expected to yield any but the most superficial knowledge of the universals of language. Similarly, if the general theory does not provide any evaluation procedure which may be applied to individual grammars, it provides no reason for choosing any one description over the indefinite set of possibilities which will always be compatible with any amount of observed data.[9]

To sum up then, general linguistic theory must contain an explicit description of all those formal and empirical constraints to be imposed on individual grammars.

CHAPTER 2: PHRASE MARKERS

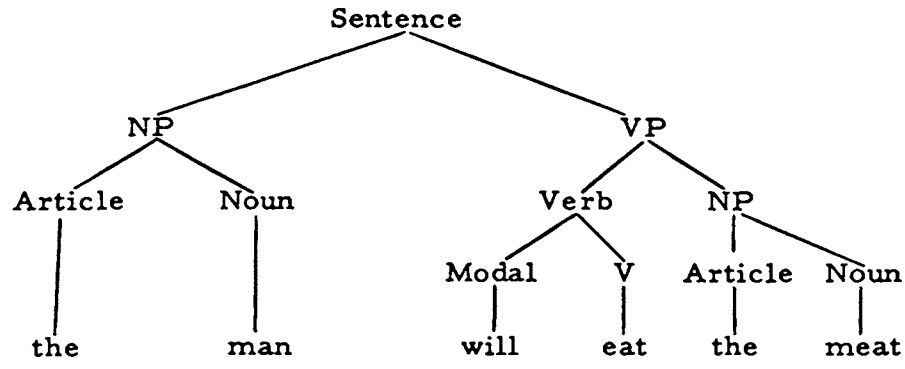
Given any sentence in a natural language represented as a string of discrete entities, say morphemes or words, certainly almost all linguists are agreed that at least part of the structure of such a sentence can be represented in the form of a hierarchical categorization of these discrete elements. Suppose the sentence is English (1) The man will eat the meat. This might be categorized as:

The man-will eat the meat
The man-will eat-the meat
The-man-will-eat-the-meat

That is, the sentence as a whole consists of two parts, The man and will eat the meat. The man in turn consists of two parts, The and man; will eat the meat consists of the parts will eat and the meat. These last segments consist respectively of the parts the, meat, will, and eat.

Furthermore, it is clear that most linguists would agree that such a categorization is inadequate unless it is associated with a labelling. Thus while the unlabelled categorization of (1) would be as above, the labelled bracketing would be something like:

Diagram 1



The labelling is necessary, if for no other reason, in order to represent correctly the similarities and differences of elements in different sentences and different parts of sentences. Thus given (2) A boy likes the girl, we not only must show that this has a hierarchical categorization like (1) but that in (2) a boy is an element of the same type as the man, likes is an element of the same type as will eat, will is not an element of the same type as meat, etc. This information is provided by giving the natural labelled bracketing for (2) and so on for an endless number of analogous and more complicated cases.

It should be noted that labelled branching diagrams like Diagram 1 have a number of equivalents such as labelled box diagrams, labelled parenthesizations,[10] etc. Following Chomsky,[11] I shall call such representations as Diagram 1 'Phrase Markers' (P-markers). The fundamental notion in the P-marker is that certain strings of elements are related to certain other single elements by the relation 'is a (member of the category)'. This relation is represented by the fact that some strings of elements are traceable back to a single node. Elements traceable back to a node A are said to be an A, or to be constituents of A. A is said to dominate such strings of elements. Thus in Diagram 1, (NP VP) is a Sentence, (Article Noun) is an NP, (Verb NP) is a VP, (Modal V Article Noun) is a VP, man will) is no constituent whatever, (Article Noun) is not a Sentence, etc. Thus P-markers appear to represent a straightforward characterization of the ideas involved in the linguistic notions of hierarchical categorization or constituency.

The conceptions represented by P-markers are really quite traditional. The earlier grammatical analysis which spoke of parsing a sentence, etc. utilized in essence just these ideas. For example, on the highest level the sentence was not only divided into two parts, but these were named 'Subject' and 'Predicate'. Similarly, on a lower level each word was assigned to one or more labelled constituents called 'parts of speech', etc.

In one way or another there is today fairly good agreement that P-markers are a correct representation of at least part of the syntactic structure of individual sentences.[12] This agreement is, of course, accompanied by many disagreements. In particular, there is controversy over how the word fits into P-markers, that is, over the place of the word in the constituent hierarchy; there is disagreement over the number of elements which may branch from any given node, especially over whether this may be more than two; and for any given sentence there will very likely be many disagreements

over what particular P-marker it should have. Abstracting from these disagreements, we may say, however, that few linguists would dispute the fact that each sentence may be correctly represented with at least one P-marker of some sort.

This fundamental agreement immediately raises a question of basic importance. Certainly we may assume that each mature speaker knows his language. If, then, sentences have structure representable by P-markers, it follows that in some sense the speaker knows these structures. But since each sentence has at least one P-marker and since there are an infinite number of sentences,[15] it follows that the speaker knows an infinite number of P-markers. How, then, is this knowledge represented in an organism with a finite storage capacity? The speaker can obviously not store such an infinite set of P-markers as such. It follows that the speaker must have internalized some general or recursive rules which can in some way represent the infinite set of P-markers. Hence, if a grammar is to be a correct explication of the speaker's implicit knowledge of his language, the grammatical description of a language cannot be a mere list of P-markers or the elements these contain. Instead it must be a set of recursive rules which can, at the very least, characterize an infinite set of P-markers.

General grammatical theory must thus concern itself with the fundamental problem of the kinds of rules which can assign P-markers to infinite sets of strings of morphemes (more generally, with the question of the kinds of device which can assign correct SD to infinite sets of sentences, even if SD are not fully equated with P-markers). In short, the general agreement that sentences have P-markers serves as a direct argument for most of the conditions assumed in Chapter 1 for grammars and linguistic theory.

Modern linguistics has, for the most part, ignored the question of the kinds of finite mechanisms the speaker might learn which would permit the assignment of SD, in particular P-markers, to infinite sets of sentences. Failure on the part of linguistic theory to describe the machinery which can account for P-marker assignment amounts, however, to a failure in principle to describe even those aspects of grammatical structure upon which nearly all linguists are agreed. The fundamental aspect of any theory of grammar which is concerned with such notions as constituent, construction, tagmeme, etc. must be the devices it makes available for the assignment of structure of this sort, in short the kinds of explicit structure-assigning grammatical rules it countenances.

CHAPTER 3: REWRITING SYSTEMS

From the formal point of view a language may be considered an infinite set of strings of elements of a finite vocabulary. Each string or sentence is of finite length, the number being infinite because there is no longest string. As has been briefly argued above, the problem of grammar construction is then minimally to provide a finite or recursive enumeration of just the set of sentences, that is, each sentence must be enumerated with its SD and no nonsentence must be enumerated.

In following discussions of formal features of grammars I shall utilize the following conventions. Early capital letters A, B, etc. will stand for arbitrary single nonterminal symbols; late capital letters X, Y, etc. will stand for arbitrary strings of symbols, terminal or nonterminal; small letters will stand for terminal symbols with early and late distinguished as above. The terms 'terminal' and 'nonterminal' will be defined precisely below. Roughly, terminals are those which are not affected by any rules.

PSG, as discussed by Chomsky, are special classes of rewriting system where this refers to sets of rules defined exclusively on strings of symbols related by the operation of concatenation. Suppose we consider a finite set of rules of the form $U \rightarrow W$, where U and W are of finite length and where W may be the identity element I ($IX = XI = X$), and the arrow is to be interpreted as the instruction 'is to be rewritten as'.[14] Rules meeting no further requirements may be called unrestricted rewriting systems.[15]

It is known that any set of strings which can be effectively enumerated at all (technically, any recursively enumerable set) can be enumerated by an unrestricted rewriting system.[16] That is, if a language has a grammar at all, it has a grammar of this highly unstructured sort. However, such systems are of no linguistic interest for two reasons. First, there is no natural way of uniformly associating an appropriate SD with each sentence enumerated by an unrestricted rewriting system. In particular, there is no way of associating a P-marker with each sentence. Second, unrestricted rewriting systems make only the absolutely weakest possible claims about language. To say only that a language has an unrestricted rewrite grammar is merely to say that the language consists of a set

of sentences which can be described. There is no point whatever in actually constructing such grammars since no insight into language may be gained by actually constructing devices whose existence for particular languages is guaranteed in advance and whose abstract character is already known.

If rewriting systems are to serve for grammatical purposes, it is then necessary to impose further conditions on rewrite rules.

Condition (1) If U → W, then W is at least as long as U, that is, contains at least as many symbols.

Condition (1), which amounts to the elimination of deletions or mappings into the identity element I, imposes an essential limitation on the generative power of rewrite systems. Every set of strings which can be enumerated by rules meeting Condition (1) is recursive, that is, both it and its complement are recursively enumerable.[17] The set of recursive sets is only a subset of the recursively enumerable sets generated by unrestricted rewriting systems. Furthermore, grammars meeting Condition (1) cannot generate all recursive sets. However, even though Condition (1) imposes an essential restriction on generative power, it still does not ensure the assignment of a reasonable SD. That is, there is still no nonarbitrary, uniform way of mechanically associating P-markers with the output of rewrite rules meeting only Condition (1).

It was Chomsky's important discovery that to insure this one must also impose at least the following restrictions on rewrite rules:

Condition (2) If U → W, then:
- a. U = XAY and W = XZY
- b. Z is not null, that is, is not the identity element I (This, of course, follows directly from Condition (1) but is repeated here for ease of reference.).
- c. Z is not identical to A.

Condition (2) insists that each rewrite rule replace a single symbol by a nonnull string distinct from the original. Rewriting systems meeting Condition (2) were called 'PSG' and studied by Chomsky in a number of publications.[18] Restrictions (2) a-c permit the mechanical assignment of P-markers to derived strings of morphemes. That is, there is an intuition-free way of mapping any derivation based on such rules into a P-marker. To make this clearer let us consider the notion of 'derivation' in greater detail.

PSG are, of course, based on a finite vocabulary of elements. Among these are a designated initial symbol S, standing for 'Sentence', and a designated boundary symbol #. Each derivation begins with the initial string #S#. A derivation is a sequence of lines beginning with #S#, each obtained from the preceding by one application of one and only one rule to a single element. In running through the rules one finally produces a line which cannot be further altered by any of the rules because none of its symbols appear on the left hand side of any rule. Such a string is called a terminal string and its symbols terminal symbols. Other symbols are nonterminal. In linguistic terms the natural interpretation of the terminal vocabulary is as morphemes, the nonterminal vocabulary as higher order constituents.

Derivations based on rules meeting conditions (2)a-c automatically formalize much of the structure involved in the ideas of grammatical category, parsing, or immediate constituent analysis because there is a relatively simple mechanical procedure for associating the formal analogue of these notions, that is P-markers, with such derivations. By definition each successive line of a derivation is formed by the application of one and only one rule. Condition (2)a guarantees that each successive line in a derivation differs from its predecessor by the replacement of only one symbol. Condition (2)b guarantees that the replacing expression is not shorter than the original. Condition (2)c prevents repetition of identical lines. Therefore, there is a finite way of representing every derivation in the form of a rooted, labelled tree or any of its equivalents by writing down the lines of the derivation with the left-most string at the top and each successively right element successively lower. Then, beginning at the top, elements are connected by lines to their corresponding identities in the next lowest line or to the strings which have replaced them.[19]

Thus, suppose there are rules:

R1 S → A B
R2 A → c d
R3 B → e f g

If these rules are unordered, there are two possible derivations (# omitted in this and all succeeding cases):

(i)
S, AB, cdB, cdefg

(ii)
S, AB, Aefg, cdefg

Each of these derivations generates the single terminal string cdefg, but the technique of tree construction just described associates with them respectively the distinct representations:

Diagram 2

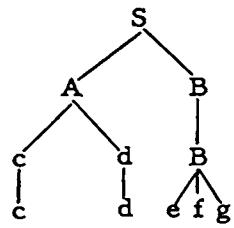
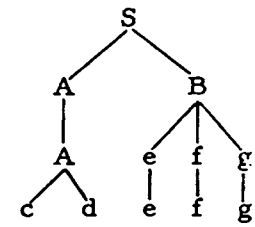
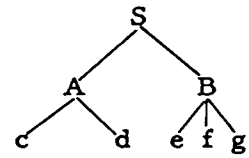


Diagram 3



However, part of the algorithm of P-marker assignment is the elimination of all but the highest elements in branches of the form $\overset{X}{\underset{X}{\vdots}}$ where all elements are identical. Diagrams 2 and 3 thus both reduce to the P-marker:

Diagram 4



A P-marker is thus a representation of an equivalence class of derivations, that is, derivations differing only in the order in which the elements have been expanded. Put differently, a P-marker is equivalent to the set of all strings which appear in a full class of equivalent derivations (notice that a derivation is not a set of strings but a sequence, that is, an ordered set). A great advantage of taking a P-marker as the set of strings which appear in the full class of derivations is that the notion 'is a', which is fundamental to constituent analysis, can be simply and directly defined in terms of substitution. That is, given this account of P-marker, the string Z is an A, just in case the P-marker contains two strings differing only in that one has Z where the other has A.[20]

Conditions (2)a-c are quite crucial if P-markers are to be assigned correctly by the above method which is the only one ever given. To show this it is sufficient to note that there is no nonarbitrary way of associating a correct labelled tree with derivations (iii) - (v) which fail conditions (2)a-c respectively:

(iii)	(iv)	(v)
S, ABC, ADEF, . . .	S, AB, ACDE, ACD, . . .	S, AB, AB, ACD, . . .

In (iii) one cannot tell which elements to attach D, E, and F to. In (iv) there is no nonarbitrary way to connect the third and fourth lines. In (v) the procedure of P-marker construction yields a tree but this would be identical with that for the quite different derivation containing no repetition. Hence, if the claim that a labelled tree represents the set of strings in an equivalence class of derivations is to be maintained, expansions of A into A cannot be allowed because these do not affect the set of strings in the full class.[21] Thus fundamental to the idea of P-marker assignment, and hence the linguistic meaningfulness of PSG rules, is the fact that each rule must expand a single symbol into a nonnull string distinct from the original and that each line of a derivation must result from one and only one rule application.

Notice that rules of the PSG type plus the algorithm of tree construction not only give a formal rendering of the notion of labelled, hierarchical categorization, that is, a precise characterization of one type of SD, but they also provide a uniform, mechanical method of assigning such structures to infinite sets of sentences. The former task has received a good deal of attention in modern linguistic work, but questions about the latter have hardly even been raised.

It is known, incidentally, that imposing Condition (2) on sets of rules meeting Condition (1) does not affect the sets of strings which may be generated.[22] Since this restriction does, however, radically effect the possibility of associating SD with the enumerated output, this is a good illustration of the fact that grammatical systems may be (weakly) equivalent in generating the same class of languages but not (strongly) equivalent in providing the same SD for the sentences of these languages. The notions of strong equivalence and generative power in terms of SD are obviously the ultimately significant ideas for linguistics.

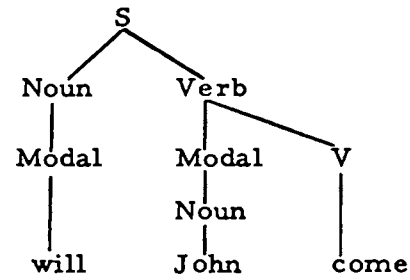
The features discussed thus far, Conditions (2)a-c, are all those originally taken to define the notion 'PSG'. However, as Chomsky pointed out,[23] this explication of the ideas of phrase structure must be considered inadequate. While Conditions (2)a-c are necessary for the correct assignment of P-markers, they are not sufficient. For within systems of rules meeting (2)a-c it is possible to achieve permutations. And the P-markers which result from permutations cannot correctly represent language structure because the asymmetry of the relation 'is a(n)' is not preserved. That is, this relation must hold between, say, Noun and NP but not between NP and Noun. But this is not in fact guaranteed by (2)a-c. Thus, suppose

sentences like John will come and will John come are to be derived by ordered rules which are sketchily:

R4 S → Noun Verb
R5 Verb → Modal V
R6 Noun → Modal in ___ Modal[24]
R7 Modal → Noun in Modal ___
R8 Noun → John
R9 Modal → will
R10 V → come

Then the latter sentence will receive the P-marker:

Diagram 5



Obviously, however, the categorization in Diagram 5 is linguistically absurd. The description asserts that John is a Modal and that will is a Noun. But this is to claim that intuitively John is similar to may, will is similar to boy, cook, meat, etc. This set of claimed similarities is simply false and no grammar which makes such claims can be accepted. Note that the absurdities involved in such claims are not particular to this example but will be a general feature of the description provided whenever sentences are derived by permutation within a PSG since, as far as I know, it will never correctly be the case that A is a B and B is an A. But there is nothing in the formulation of phrase structure given thus far which predicts that such descriptions are unacceptable or in any way different from ordinary intuitively correct analyses. Indeed, given the otherwise correct evaluation procedure which measures generalization or simplicity by counting symbols in the rules under certain notational coalescings, these incorrect solutions would in fact be the simplest and thus the formally motivated ones for many sets of sentences. It is thus clear that the characterization of phrase structure must include another condition to exclude the possibility of permutations.[25]

Condition (3) If XAY → XBY, there are no rules of the form:

$$U_1BW_1 \rightarrow U_1C_1W_1$$
$$U_2C_1W_2 \rightarrow U_2C_2W_2$$
$$\ldots\ldots\ldots\ldots$$
$$U_nC_{n-1}W_n \rightarrow U_nC_nW_n$$

where $C_n = A$ (Note that $C_1 \ldots C_n$ are single symbols.)

That is, if A is expanded into B in some context, then there are no contexts such that B is expanded into A or into anything which is expanded into A; or into anything which is expanded into anything which is expanded into A, etc. This condition eliminates the possibility of permutations (note that it excludes R7 above) and, I think, guarantees the correctness of P-marker assignment insofar as this is possible within rewrite systems at all. Henceforth, I shall use the term 'PSG' to refer only to systems meeting Condition (3) as well as (2). I wish to argue in this study that a large class of proposed systems of descriptive grammar are, with certain exceptions, versions of the theory of PSG in this sense.

Before passing to such considerations, however, it is necessary to mention one further constraint which provides a division of all PSG into two important types.

Condition (4) If XAY → XZY, X and Y are null. This requires that if a symbol A is expanded into a string Z, then it is expanded into Z in every sequence in which it occurs. Systems all of whose rules meet Condition (4) will be called CF-PSG (context-free), those whose rules do not all meet this condition will be called CS-PSG (context-sensitive). We shall see that this condition plays an important role in contemporary grammar and later we shall have some very crucial empirical results to present which determine the truth value of this constraint.

It should be emphasized that the theory of PSG is not without technical flaws even given Condition (3). Chief among these is the fact that for CS-PSG the uniqueness of P-marker assignment is undecidable. That is, given an arbitrary set of CS-PSG rules it is provably impossible to tell whether there is one and only one P-marker assignable for each derivation. Thus CS-PSG fail to meet the minimal requirement of mechanical assignment of structure to generated sentences.

This nonuniqueness comes about through the possibility of successive lines in derivations like ABC, ADBC. Here there is no way for the procedure of P-marker construction to determine whether A has expanded into A D, or B into D B. For CF-PSG this nonuniqueness is also possible, but in these systems, given a set of rules, it is always possible to tell whether or not it occurs. Thus, for CF-PSG uniqueness of P-marker assignment is decidable. Put differently, since the notion of 'grammar' requires uniqueness of structure assignment, given any set of CF-PSG rules we can tell whether or not it is a grammar. The question of uniqueness can obviously be made decidable for CS-PSG by requiring that the set of derivations be finite. There are a number of motivations for such a restriction. But this would immediately mean that PSG-rules must be supplemented by more powerful devices, as in a TG.

Independently of a restriction of PSG derivations to a finite number, the uniqueness of P-marker assignment can be guaranteed by strengthening Condition (2)c to: If $XAY \rightarrow XZY$

$$\text{Condition (2)c}' \quad Z \neq (U)\ A\ (W).$$

This condition prevents not only the expansion of A into A but also the expansion of A into any string containing A. This prevents successive lines like those illustrated in the previous paragraph which cause an indeterminancy in P-marker assignment, given a derivation.

It is interesting to note that this very natural condition prevents the only natural PSG description of both coordinate constructions and endocentric derivation. The former requires rules like NP → NP and NP; the latter those like NP → Adjective NP. This is an important conclusion because there is independent evidence in terms of both simplicity and correctness of SD to show that these processes of sentence formation are transformational. PSG description of coordinate constructions and endocentric derivation always leads to incorrect P-markers (see the discussion in Chapter 4: section III) by arbitrarily imposing too much structure and thus failing to distinguish the coordinate P-markers which such constructions should have from the subordinate P-markers of embedding constructions. Condition (2)c' suggests that the transformational character of these constructions may be related to the need to insure uniqueness of P-marker assignment. And although Condition (2)c' does not have this property, it is possible that there is some correct formal condition on PSG rules which makes it literally impossible to describe coordinate constructions in such terms (even given Condition (2)c', it is still possible to describe coordination by using ad hoc rules like:

Nominal → NP and NP; NP → Nominal). This discussion reveals some of the ways in which study of the abstract formal properties of grammatical theories may relate to the empirical correctness of linguistic descriptions.

Despite the virtues of Condition (2)c' and the fact that a correct explication of the notions of phrase structure requires it or its equivalent, we shall not impose this condition on rewrite systems in this monograph. PSG, in the sense of the present study, are thus subject only to the weaker Condition (2)c.

CHAPTER 4: SOME PROPOSED MODELS OF GRAMMATICAL DESCRIPTION

I. Remarks

In the present chapter we present an explication of various grammatical proposals which have been made, attempting to show how they may be interpreted in terms of generative rewrite rules, in particular in terms of PSG. It should be emphasized that we do not claim to prove that the views discussed are equivalent to any formal system. This is for the most part not possible because the ideas to be considered are unformalized. It is difficult, however, to determine just what inferences may be drawn from imprecise theories. This is just the motivation for formalization. Unformalized notions may conceal several alternative conceptions, hidden contradictions, unspecified questions of real theoretical interest, etc. The task of discussing modern grammatical ideas in the light of formal generative rules is made doubly difficult because, with few exceptions, these ideas are discussed from a nongenerative point of view. The major emphasis is on types of SD, the elements these contain, and the way that these may be derived or discovered. Nonetheless, I think that in most cases it will be possible to reach a good degree of determinateness with respect to a formal interpretation of most features of the class of grammatical views chosen for discussion; for the rest we shall be able to pinpoint various gaps, unspecified properties, etc., that is, we shall be able to show just where the ideas considered are not explicit on points of linguistic interest.

II. Bloch's Japanese Syntax

Bloch's study of Japanese Syntax[26] presents a clear instance of a descriptive grammatical work which has, for the most part, a straightforward interpretation in terms of PSG. Bloch first distinguishes different sentence types, most important among these being the Major Sentence Type. This division can be naturally accounted for in PSG terms by rules such as:

$$\text{R11 Sentence} \rightarrow \left\{ \begin{array}{l} \text{Major Sentence} \\ \text{Exclamations} \\ \text{etc.} \end{array} \right\}$$

Such rules then indicate that Sentence has several different expansions, each of which provides this highest order constituent with a single lower level constituent.

Bloch then systematically discusses the elements on the right hand side of R11, dividing them successively into more and more differentiated sets of exhaustive, nonoverlapping, continuous elements, their 'immediate constituents' on each level. Each statement of constituency indicates the order of constituents and their names. Thus a typical statement is:[27]

> 'An adjective phrase consists of an adjective as nucleus preceded without pause by an inflected expression in participial form'.

Clearly, statements of this type can be translated into PSG rules of the form:

R12 Adjectival Phrase → Inflected Expression Adjective

or perhaps into pairs of rules like R12 and:

R13 Inflected Expression → Participial

and so on for all other of Bloch's statements of immediate constituency insofar as constituents are continuous. Perhaps the only notion not reconstructed in the PSG rules is that of 'nucleus'. We return to this idea of 'nucleus' or 'head' below in our discussions of some of Harris' formulations.

It would appear that in order to provide the sentence description given by Bloch, at least some of the rules would have to contain contexts. Thus we are told that both final and nonfinal clauses contain predicate elements, but that the predicates of final and nonfinal clauses are different in form. This type of structural information can only be provided if the Predicate constituent is expanded differently according to context, that is, one way in final clauses, another way in nonfinals. There are also other cases where Bloch's description would seem to require interpretation in terms of CS-rules. We return below to a consideration of conditions under which a context sensitive interpretation is required and to a discussion of the possibility of replacing CS-rules with CF ones.

A feature of Bloch's description which deserves mention is the way the labelling of constituents is treated. This is an important matter because, as noted above, the notion of P-marker involves not only a hierarchical bracketing, but also an associated labelling. Bloch names his constituents for the most part, using such terms as 'Final Clause', 'predicate', 'adjective phrase', etc. But it is at least possible that these labels were meant as nothing more than a shorthand way of referring to data. In claiming a PSG interpretation for most of Bloch's description, I assume that there is as much a truth claim involved in the labelling as in the bracketing. The question of the place of labelling in constituent structure is important since it has been a source of some confusion and, I think, several pseudo problems. We shall see below that one aspect of the so-called 'tagmemic' version of syntactic description is a very strong and, I think, entirely correct emphasis on the need to consider a labelling of constituents as a part of language.

If there were no other aspects of Bloch's description besides those discussed above, all of Bloch's description could readily be interpreted in terms of CS-PSG. However, Bloch allows for the existence of some 'discontinuous constituents' and structure of this type cannot, of course, be assigned by PSG. It follows that to the extent that Bloch allows such structure, his description is not completely interpretable in terms of PSG. As with many other writers, including those dealt with below, Bloch does not raise the question of how discontinuous structure is to be assigned. More generally, there is no discussion of how linguistic structure in general is to be assigned and this follows naturally from the great emphasis placed on techniques for discovering the grammar and a corresponding lack of attention to the actual character of the description which results. However, all those features of sentences discussed by Bloch with the exception of discontinuities can be characterized by PSG rules. Since several other writers considered below allow discontinuities, we shall discuss this topic in a separate chapter.

Bloch's study is historically the earliest one we shall consider and one of the earliest formulations of a syntactic description within the modern American tradition. It is important to note that in both clarity of presentation and empirical adequacy of the range of ideas used, it has not been exceeded by later formulations. Indeed, as we shall see below, as descriptive approaches become historically farther removed from Bloch's description, they tend to become if anything less clear and less adequate.

III. Wells' Immediate Constituent Approach

Rulon Wells' influential article 'Immediate Constituents'[28] is often taken as the highest or classical statement of the immediate constituent approach in American linguistics. Most of Wells' discussion is concerned with the question of a methodology which might be used to determine the proper constituent analysis of any presented sentence. As is typical in taxonomic studies, there is little attention paid to the character of grammatical descriptions as such. There are no explicit statements about the possible form of linguistic rules or the way these assign structure to infinite sets of sentences. Since we are not concerned with methodology as such in this study, that part of Wells' article which concerns us is related chiefly to the types of SD he considers.

The term 'IC analysis' occurs frequently in Wells' work. He uses this to mean an unlabelled hierarchical bracketing of a string of morphemes into segments, for the most part continuous. These are typically indicated by varying numbers of vertical lines. Thus an example would be: paint | the || new ||| store. The most major or highest order break is indicated by a single line, successively lower order breaks or 'cuts' by greater numbers of lines. This analysis then corresponds to a tree representation without labels. The equivalent of labelling is introduced, as in the case of Hockett's writings considered below, by the use of the notion 'construction'. Wells' defines this notion in a classificatory way which does not really concern us. Indeed he admits in effect that the definition is untenable.[29] However, the fact that 'construction' amounts to a reference to a labelling of constituents is shown in the following:[30]

> 'However, the two meanings of old men and women are most readily accounted for in the following way. In the meaning 'women and old men', the sequence belongs to that construction (noun or noun-phrase + and + noun or noun-phrase) which has the meaning of conjunction; the first noun-phrase belongs to the construction modifier + noun or noun-phrase. But in the meaning 'old men and old women', the sequence belongs to the construction modifier + noun or noun-phrase; the noun-phrase in turn belongs to the construction noun or noun-phrase + and + noun or noun-phrase.'

Clearly this is to say, in effect, that the two different interpretations of the phrase can be represented by the two different P-markers:

Diagram 6

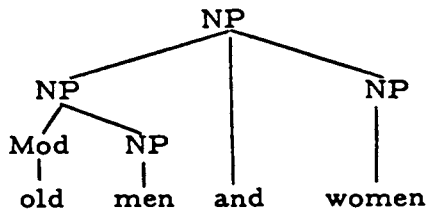
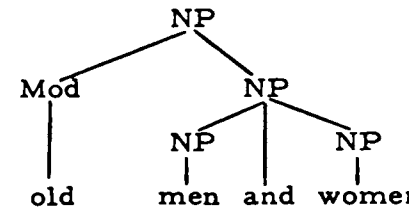


Diagram 7



It thus seems clear that taken together Wells' notions of IC analysis (insofar as this is continuous) and construction are equivalent to labelled bracketings or, in other words, to P-markers.[31]

The typical situation for Wells is that in which the degree of branching is two, that is, where a constituent dominates two other constituents. Indeed, unlike Bloch, he does not allow a situation where a constituent has a single IC.[32] Wells thus adopts, in effect, the restriction suggested by Parikh which we mentioned earlier. This proveably eliminates permutations. For other views which we consider, including Bloch's, our assumption that Condition (3), or its equivalent with respect to eliminating permutations, is met is based on the fact that absurd representations of constituent structure of the type produced by permutations would never be allowed. As noted when Parikh's restriction was first mentioned, it is too strong. Wells discusses the fact that no constituent can, for him, have one IC in terms of an example which Bloch handled by allowing just this. The solution which Wells suggests is quite clearly inadequate and possibly inconsistent. At any rate, there are many other examples where rules of the form XAY → XBY are required. It is interesting that Wells offers no support for his refusal to allow one IC. This, then, is the first instance in which the descriptive apparatus becomes less adequate with respect to Bloch's formulation.

All binary structures may, of course, be enumerated by PSG, indeed by the special class of PSG in which no rule has more than two symbols on the right hand side. But Wells allows for richer SD. Like Bloch, he allows discontinuous constituents and thus to that extent his system lies outside of PSG interpretation. And Wells also recognizes that in some cases it is necessary to recognize multiple ICs, that is, cases where a single constituent dominates three or more elements.

He proposes a principle to determine where the above is necessary.[33]

> 'Given a constitute[34] consisting of three continuous sequences A, B, and C, then, if no reason can be found for analyzing it as AB][C rather than A][BC, or as A][BC rather than AB][C, it is to be analyzed into three correlative ICs, A][B][C. Similarly, four ICs may be recognized when no analysis into two and no analysis into three ICs is recommended, and so on.'

The recognition of multiple ICs does not in itself place the descriptive system beyond PSG interpretation in the way, for example, that discontinuities do. For multiple ICs can be handled by PSG rules. For example, quaternary analyses can be described by rules of the form:

R14 A → B C D E

quinternary analyses by:

R15 A → G H I J K

etc. But note that in any PSG there must be some analysis with a branching lower than some fixed finite n since there are a finite number of rules each with a finite number of symbols on the right hand side.

But if we take seriously Wells' 'principle' that natural languages contain constructions which should have n-ary branchings, where n is the number of constituents they contain, we can find cases where this number is unlimited. These are, of course, the so-called 'co-ordinate' constructions such as those with and in English. A PSG description of such constructions must contain a separate rule for each length and thus, impossibly, an infinite number of rules, or else it must contain rules like:

R16 Nominal → Nominal and Nominal

However, rules like R16 assign P-markers like Diagram 8 instead of the correct Diagram 9:

Diagram 8

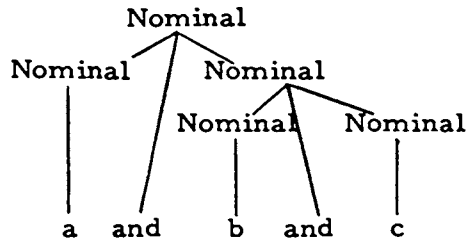
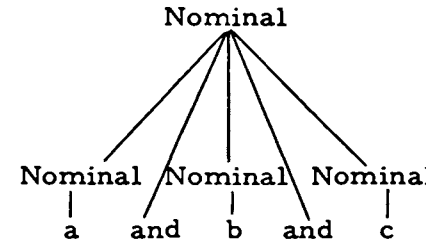


Diagram 9



Thus, as shown by Lees,[35] PSG necessarily assign the wrong P-markers to coordinate constructions through arbitrary assignment of excess structure. Furthermore, the rules needed by PSG to generate coordinate strings of boundless length violate Condition (2)c', needed to insure uniqueness of P-marker assignment in a natural way.

Insofar as the fact that coordinate constructions require branching greater than any fixed n may be derived from Wells' principle of multiple cuts, it follows that Wells' system recognizes SD that are beyond PSG enumeration in a really crucial way. But since he did not explicitly develop his principle to this conclusion and since he did not raise the issue of the kinds of device which could enumerate the types of SD he recognized, this insight on Wells' part appears to have had no effect on the development of grammatical theory.[36] This is a rather striking case of the fact that even entirely correct observations about language make little contribution to our understanding of grammar if they are not framed in terms of an explicit conception of the nature of grammatical rules.

I am unable to determine from Wells' article whether or not the PSG interpretation of his IC approach would require contexts in the rules. This can only be determined if there is at least some characterization of the possible rules or else if there is an explicit discussion of conditions under which two sequences in different sentences or different parts of the same sentence are 'the same constituent'. As noted earlier, the primary emphasis in Wells' work is on unlabelled bracketings and the notion of 'same' constituent, that is, those sequences with the same label, is handled under the notion of construction. While we cannot determine whether or not Wells' 1947 views countenanced contexts in rules, it should be noted that he was aware of the fact that constituents could not be set up in such a way that all instances of a single constituent shared all environments,[37] or, as he put it, so that they were all absolute equivalents. As we note later in discussing Hockett's 'constructional

grammar', imposing this condition amounts to the assumption that all rules are CF. However, not imposing this condition does not mean that all the rules are necessarily not CF and thus the indeterminateness of Wells' IC ideas with respect to contexts in rules remains.

IV. Harris' Morpheme Class Substitution System

In an article and a book, written at nearly the same time but published several years apart,[38] Harris discusses at length a method of obtaining the syntactic description of a language by means of substitutibility operations. We are interested here not in these operations, but rather in the grammatical description which may be expected to result from them.

Harris assumes that every sentence is represented by a sequence of morpheme classes containing mutually substitutible morphemes. The syntactic description is then determined by finding sequences of morpheme classes which are substitutible for individual classes:[39]

> 'We equate any two sequences of classes if one of them is substitutible for the other in all utterances in which either occurs.' 'More generally, given the sequence of morpheme classes X occurring in the range of utterance environments M, we find all sequences Y, Z, etc. which occur in precisely that range and write Y = X, Z = X, etc.'

And also:[40]

> 'Equations will be used to indicate substitutibility. BC = A will mean that the sequence consisting of a morpheme of class B followed by a morpheme of class C can be substituted for a single morpheme of class A.' 'When we want to say that A substitutes for B only if C follows, we shall write AC = BC.'

Note that if we are to interpret the output of Harris' methods in generative terms, interpreting these substitutibility equations as grammatical rules, we must assume that there are a finite number of such equations from which all possible substitutions are 'derivable'. Up to this point it would appear that Harris' equations can be simply indicated by PSG-rules. Equations of the form BC = A can be given as:

R17 A → B C[41]

Rules of the form AC = BC or, more generally, ADEC = BC can be interpreted as CS-rules of the form:

R18 B → A in ___ C
R19 B → A D E in ___ C

It thus is clear that Harris permits rules with at least contexts on one side of the element to be expanded. Thus he gives the equation:[42] $A^2A^2N^1 = A^2\ N^1$. This says that a sequence of two adjectives can substitute for a single adjective if a noun follows. The restriction is presumably necessary because of the nonoccurrence of *She is big, brown by substituting from She is big.

Harris introduces raised superscripts in his equations at an early point, as in the above example. He says:[43]

> 'Sequences of morpheme classes which are found to be substitutible in virtually all environments for some single morpheme class, will be equated to that morpheme class: AN = N means that good boy, for example, can be substituted for man anywhere. If we write DA = A . . . DA can be substituted for A wherever A appears. There is nothing to prevent us from substituting DA for A even in the equation DA = A. We would then obtain DDA = A: really quite old for old. If, however, it proves impossible to substitute the equivalents of a symbol for that symbol in some of its occurrences, we distinguish those occurrences by giving the symbol a distinctive raised number. For instance, N - s = N: paper+ - s = paper; and papers can be substituted for paper in most environments. However, we cannot substitute N - s for the first N in this very equation: we cannot substitute papers for the first paper and then add - s again (papers+ - s), as this equation would seem to indicate. We therefore write N^1 - s = N^2 and state that wherever N^2 occurs we can substitute for it any N^1 or another N^2 while for N^1 we can only substitute any member of N^1 (never N^2). Then it becomes impossible to construct a sequence papers+ - s since papers is N^2 and - s is added only to N^1.'

In view of the fact that, as we have seen, Harris in effect already allows himself context restrictions without bothering about superscripts, it is interesting to ask why these are introduced and just what they imply. We can introduce this question by noting that there are, a priori, two different ways of interpreting Harris formulas with superscripts in PSG terms. On the one hand, each of

$X^1, X^2, \ldots X^n$ can be interpreted as distinct constituents from one another. Thus $TN^2 = N^3$ might be reformulated as:

R20 NP → T N

On the other hand, Harris insists[44] that the same symbols with different superscripts are the same class and contain the same single morpheme members. If this is taken seriously, it must be assumed that there is some more direct relation between N^2 and N^3 than indicated by R20. In fact, note that there is a type of information indicated by Harris' formula not given in R20, namely, that in the sequence TN^2, N^2 is the head. I think there is good reason to believe that much of the motivation for introducing superscripts is to reconstruct this notion.[45] If this can be considered successful, we would be forced to reject a PSG interpretation for Harris equations involving superscripts and say that these involve some other type of (as yet not formally characterized) system. But this attempt to reconstruct the notion 'head' or 'nucleus' cannot be considered successful.

In a formula like $TN^2 = N^3$ Harris explicitly says that on the left hand side N^2 also stands for N^1.[46] But now if N^3 is to have any function in the system, it must occur on the left hand side of some formula and indeed it does ($N^3PN^4 = N^3$).[47] But this means that in effect $TN^2 = N^3$ says that the distribution of N^2 is included in that of TN^2. But the analogous claim for all similar formulas is obviously in general not true. For it to be true for arbitrary morpheme classes it would have to be the case that all constructions are 'pseudoendocentric', that is, such that the head is always substitutible for the expansion even if the expansion is not fully substitutible for the head. Fully endocentric constructions are then a subset of the pseudoendocentric ones in which there is mutual substitutibility in both directions. But obviously all constructions are not of the pseudoendocentric type. Thus for the example discussed, TN^2 cannot be replaced by N^2 everywhere, for example, not in Some of [the men] are nice since there is no *Some of men are nice. Thus Harris' use of the superscripts to indicate the head of a construction cannot be taken literally. The only interpretation of his formulas which can avoid the false claim that all constructions are pseudoendocentric is the one in which the superscripted symbols simply indicate different constituents.

It should be noted that in basing his constituent system on substitution for individual morpheme classes Harris makes the quite clearly false claim that all constituents in all languages are typically substitutible for a single morpheme. In languages (like Iroquoian)

where many constituents consist of sequences of bound morphemes the untenability of this claim is quickly evident. For example, in Mohawk the Verb constituent consists minimally of a pronominal prefix, a verb stem, and a suffix. There is no single morpheme constituent which can in general substitute for such a sequence. Thus Harris' approach would not permit the Verb constituent in Mohawk.[48]

V. Hockett's Item and Arrangement System

In his article 'Two Models of Grammatical Description', first published eight years ago, Hockett presented a succinct statement of what he called the item and arrangement (IA) version of grammatical description:[49]

> 'The tactical pattern of a language is completely covered by a set of statements of the following form (or by any set of statements which can be transformed mechanically into this form):
> (1) A list of the constructions.
> (2) Under each construction as heading,
> (2.1) Enumeration of the positions in that construction.
> (2.2) Specification of any marker for that construction.
> (2.3) For each position,
> (2.31) A list of the morphemes which occur there, and
> (2.32) A list of the constructions, composite forms belonging to which occur there.'

Earlier we were told about constructions:[50]

> 'A composite form consists of two or more immediate constituents standing in a construction and forming a constitute. Constituents and constructions recur in other composite forms. . .' 'Each IC . . . occupies a certain position in the construction.'

This conception of grammar has, I think, a fairly obvious interpretation as a PSG. The list of constructions as well as the list of positions in each can be directly provided by PSG rules. Construction here simply seems to mean the labelling of constituents. In a list of PSG rules the set of elements on the left hand side provides the list of constructions and the elements on the right hand side provide the list of positions for each. Thus a rule like:

R21 Verb Phrase → Verb Object

says that there is a Verb Phrase construction which has two positions, the first of which is filled by the form class Verb, the second by the Object construction. Verb and Object are the immediate constituents.

The term 'constitute' in Hockett's usage appears to refer to a particular sequence of constituents which is in fact a higher order constituent while abstracting from the higher order label. Thus, I suppose, the same sequence of constituents might be a constitute in several different contexts, representing a different construction in each. In a grammar containing R21, (Verb Object) is a constitute.

Hockett's (2.31) can, of course, be provided for by PSG rules like:

R22 Noun → John, boy, etc.

His (2.32) can be handled by rules like:

R23 Verb → V Infinitive

etc. This says, in Hockett's terms, that the first position in the Verb Phrase construction as characterized by R21 may be filled by composite forms of the type (V Infinitive).

Thus, insofar as all the constituents of a construction are continuous, it is clear that Hockett's IA model has a direct PSG formulation except, perhaps, for the notion of marker to which we return later in considering some of the more recent formulations of Hockett's ideas. In this work published in 1954, Hockett does not specify whether or not ICs may be discontinuous, so we shall return to this later in the chapter devoted to this topic.

We can therefore conclude that the IA conception of grammar is a PSG system meeting at least Conditions (2) and (3). It cannot be determined precisely whether Condition (4) is met, that is, whether or not contexts would be permitted in the rules. The general impression given is that such would not be permitted.

In discussing the IA system, I would not like to be thought to be doing an injustice to the item and process (IP) version given by Hockett in the same article. But, insofar as I can understand his formulation of the latter, they are equivalent; differing only terminologically with the term 'process' replacing the term 'construction'.

VI. Lamb's Stratificational Syntax

In a recent publication,[51] Lamb presents in brief an outline of a form of syntactic description. Interpretation of this approach is made difficult, however, because of the brevity and obscurity of its presentation, the fact that none of the suggested devices are utilized or illustrated with actual data, and because questions of the determination of simplicity and devices for this purpose are mixed with considerations of the actual mathematical form of the rules to be permitted.[52]

Lamb regards the syntactic or tactic description as consisting of a list of tactic rules. These have three parts:

a. Left Half
b. $\subset$ or =
c. Right Half

The left half consists essentially of any finite string of distribution classes, including the null class which Lamb writes as '1'. The right half consists of a single class symbol. Lamb allows for intersection of classes (indicated by '.') so that on the left of some rules there may be sequences like A.B, A.B.C. etc., where these refer to the intersections of the classes A and B, and A and B and C respectively. However, these complex symbols are actually just the names of individual classes, namely, the intersections. Thus in fact the right half of all tactical rules consists of a single symbol.

$\subset$ and = are the operators which connect the two halves of the rules. $\subset$ is the inclusion sign for classes. $A \subset B$ if and only if all x which are members of A are members of B. A = B if the above and its converse hold. The difference between these two operators in Lamb's rules concerns only his notational conventions for determining simplicity[53] and does not affect the deeper structural properties of the rules with which we are concerned. We can thus disregard this difference and regard all rules as having the form $B_1 B_2 \ldots B_n \subset A$, where this is to be interpreted as saying 'the class A has as a member the sequence of classes $B_1 B_2 \ldots B_n$'. But obviously, with one exception, each such rule has a CF-PSG equivalent of the form:

$$R24\ A \rightarrow B_1 B_2 \ldots B_n$$

The one exception is not without importance. It was noted that the string of classes on the left of Lamb's rules and thus on the

right of their phrase structure equivalents may include the null class. Thus his tactical rules in fact include deletions of the form:

R25 A → I

It is known[54] that deletions do not increase the weak generative power of CF-PSG rules. That is, there is no set of strings which can be enumerated by using CF-PSG rules and deletions which cannot also be enumerated by using CF-PSG rules alone. Thus rules of the form R25 provide Lamb's model with no extra generative power in the weak sense. But, as we noted above, rules of this type are incompatible with a uniform assignment of P-markers to generated strings in a non-arbitrary way by the only method known. Hence, by allowing this kind of rule within a rewrite system Lamb has in fact greatly weakened its generative power in the strong sense. That is, he prevents his rules from describing the constituent structure of sentences correctly. His claim that:[55]

> 'Using this system, the analyst arrives at distribution classes and construction rules, which describe arrangements in terms of the classes in the simplest possible terms. Questions concerning immediate constituency are automatically answered; ...'

must thus be considered false in terms of the only precise method of P-marker assignment ever given.[56]

If the apparatus discussed thus far were all that Lamb provided, his system would clearly have the weak generative power of CF-PSG but not the structure assigning properties of the latter. However, there are indications that Lamb wishes to enrichen his descriptive devices in two different ways. He introduces symbols which have associated subscripts. These, we are told, are for the purpose of describing concord. He unfortunately does not tell us how these subscripts are to enter into the rules. They are not mentioned in the section on tactic rules. We return to the question of subscripts, and more generally, to the problem of concord in a later section.

A second mechanism which would perhaps go beyond PSG in another way is also introduced by Lamb in such a way as to make it extremely difficult to determine its precise character. A dot notation is introduced, apparently to indicate permutations. Thus:[57] AB = °BA = BA°. Presumably, although this is not said explicitly, this notation is to be used to deal with discontinuities. In other words, it appears that the raised dot notation permits classes which

have the order 1 2 to be written in the order 2 1. It seems likely that this is a device to permit the permuting of directly adjacent elements. But Lamb does not specify how the dot notation enters into the tactical rules or how it may be used in actual cases. Questions of structure assignment are not raised for this device, and this is true for the rest of Lamb's system as well. I suspect that the dot notation is probably to be considered a more general though less explicit version of a device suggested by Yngve[58] in which PSG rules would be supplemented by rules of the form:

R26 A → B ... C

Given a string EAD and an application of R26 the result is to be the string EBDC. We return to a consideration of rules of this type in Chapter 5 below, which is devoted to discontinuities.

In describing stratificational grammar we have discussed only the so-called 'tactical rules'. It should be mentioned that Lamb's work also includes another type of device, the so-called 'representational rules'. I have not mentioned these because it seems that they are not intended to affect generative power, being only concerned to map one type of representation into another. If, however, these rules <u>are</u> allowed to be recursive, Lamb's system is at once simply a version of the theory of unrestricted rewriting systems since it seems that his representational rules include any possible type of rewrite operation on strings. Thus the interpretation made in this study, that the representational rules do not play a role in productive power, is the kindest possible, since it preserves Lamb's conception from the vacuity (discussed in Chapter 1) of considering unrestricted rewriting systems as a model for the structure of grammar.

We should take note of a possible argument which might be given to show that stratificational grammar is not a version of the theory of CF-PSG. One of the properties of this model is that a grammar drawn to its specifications contains not a single set of tactical rules, but a group of such sets, one for each different 'strata'. It might be suggested that this fact destroys the equivalence between stratificational grammar and CF-PSG demonstrated by considering only a single set of stratificational rules. This, however, is not the case. The number of strata is irrelevant if each contains only CF-PSG rules since the only result of recognizing more strata is the introduction of more CF-PSG rules. But the grammar G' formed by adding a finite set of CF-PSG rules to a CF-PSG grammar G is also a CF-PSG grammar. Thus the number of strata recognized by Lamb in no way affects the PSG characterization of his system.

Our discussion of Lamb's stratificational grammar can be concluded by noting that, excluding the representational rules, those portions of his descriptive apparatus which are clearly enough formulated to interpret in generative terms reduce to CF-PSG with the added disabilities of deletions.

VII. Tagmemics

In view of the importance of tagmemics today, that is, the large number of linguists who are applying it, the large number of languages to which it is being applied, the fact that these languages are, for the most part, exotic enough so that the tagmemic descriptions of them may very well be the only ones done, and the large and growing body of published literature on this view, tagmemics deserves a more thorough study than is possible here.

With few exceptions, attention is restricted here to the formulation of tagmemics given by Elson and Pickett in their introduction.[59] The major descriptive device provided by tagmemics for the grammatical description of language is the tagmemic formula. There are thirty-eight tagmemic formulas in Elson and Pickett's introduction. Of these, thirty-five have the form (in representations of tagmemic formulas our conventions for capital and small letters are abandoned):

$$A = B{:}b \quad C{:}c \quad D{:}d \ldots N{:}n$$

where this is to be interpreted as saying that the tagmeme A has as lower order representatives the string of tagmemes B:b ... N:n. Each tagmeme is thought of as a 'slot-class correlation', or, in other terms, as a 'form-function composite'. This, then, is the force of the complex symbols B:b, etc. B represents the slot name or 'function', b the class of elements which fills that slot or the form. Each of the tagmemic formulas has a single symbol on the left. These are not directly interpretable as PSG rules, however, because of the compound symbols B:b, C:c, etc. However, it is clear that these may be looked upon simply as a method of collapsing distinct phrase structure rules. That is, a formula like: A = B:b C:c may be viewed simply as a notational variant of:

R27 $A \rightarrow B\ C$
R28 $B \rightarrow b$
R29 $C \rightarrow c$

Thus each tagmemic formula of the above type can be viewed as a notational variant of a finite set of CF-PSG rules.

In view of the fact that tagmemic formulas of the type just discussed are the major descriptive tool used in tagmemics and in view of the fact that they are so easily interpreted as PSG rules, it is important to ask why the formulators of tagmemics believe that in a rather deep sense they have succeeded in formulating a new theory which is above and beyond immediate constituent analysis of the ordinary Bloch-Wells-Harris-Hockett sort. There are several reasons for this.

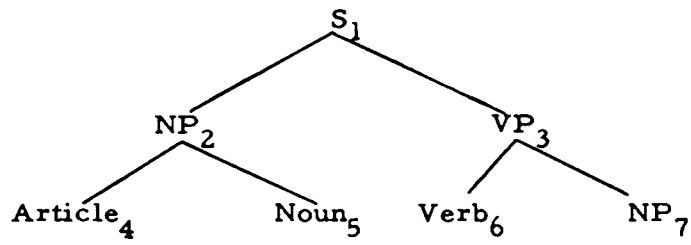
First, writers on tagmemics generally associate the term 'immediate constituent analysis' with bracketings which are binary. That is, in generative terms they equate PSG with PSG in which each rule has at most two nonterminal symbols on the right hand side. And an aspect of tagmemics which is continually emphasized is the fact that most analyses are not binary. But since, as we have seen, writers like Wells recognized multiple breaks, this is not a real theoretical departure from past American views but at best a difference in choice of analyses sentence by sentence. That is, it is not a difference which affects the type of descriptive apparatus provided by linguistic theory, but rather the particular sets of rules selected for the description of particular sentences. There are some real factual claims about language involved here (concerning perhaps statistical predominance of certain types of branching), but these do not in any way affect the PSG interpretation of tagmemic description.[60] This would only be the case if tagmemic writers had recognized not only that many branchings are multiple, but that some are infinitely multiple. But this they have not done and tagmemic formulas are incapable of describing unbounded branching correctly, that is, without arbitrary assignment of excess structure.

A second and most important reason why tagmemic writers believe their ideas to be different from previous formulations is because of their consistent, emphatic, and quite correct emphasis on the need for a labelled bracketing. The notion of tagmemic analysis inherently involves an association of labels with constituents, namely, the tagmemic names. As was noted in the discussion of the work of Bloch and Wells, the labelling of constituents in their writings was, to a certain extent, peripheral. Chief emphasis was on the division of sequences into segments. Labelling was brought in somewhat marginally by the use of the concept 'construction'. In tagmemics the notion of labelling is primary. In fact, Lees[61] took the tagmemic emphasis on form-function composites and hierarchical levels to be

simply an insistence on labelled bracketing. However, while this emphasis on labelling is important, neither it nor the relaxing of the constraint to binary branching can really be called new. Tagmemics has at best reemphasized aspects of IC analysis which have always been present, if somewhat neglected. But in the tagmeme concept there is at least one new claim. This concerns the notation B:b and its function-form, slot-filler class interpretation.

Notice that ordinary immediate constituent trees would often have a branching structure like the following:

Diagram 10



I have numbered the nodes of this subtree for ease of reference. The important formal feature which should be emphasized is that a constituent may dominate a string of constituents and each of these in turn may itself dominate a string of constituents. Thus node 1 branches into 2 and 3, 2 branches into 4 and 5, etc. However, tagmemic description does not permit this. Indeed exactly what the assumption of form-function or slot-class correlations really amounts to besides an emphasis on the labelling of the bracketings is the denial of this possibility. For a tagmemic formula of the form A = B:b C:c ... N:n asserts that A branches into B C D .. N but that the nodes corresponding to these constituents cannot themselves branch. Rather, each constituent which itself appears as a member of a string which branches from some node must be developed into a <u>single</u> symbol which is the filler class of the tagmeme denoted by the capital letter in our schematization. That is, formally tagmemics adds to Conditions (2) and (3) the condition:

> Condition (5) If $XAY \rightarrow XZY$ and $Z = B_1 B_2 \ldots B_n$ (n greater than 1), then $B_j \rightarrow U$ (where U is a single symbol, terminal or non-terminal).

Far from placing tagmemic description beyond PSG interpretation, Condition (5) restricts tagmemics to a subclass of possible PSG

grammars, a subclass whose weak generative power is identical to that of the class not meeting this condition. The recognition of more categories or constituents in no way changes the PSG characterization of the system if this can otherwise be maintained.

Note that one of the implications of Condition (5) is that there can be no rule of the form:

R30 A → Z c (Note that c is a single terminal symbol).

We discuss this further below in our treatment of agreement, showing that it causes pointless redundancies. The general effect of Condition (5) is to necessitate the recognition of roughly twice the number of nodes per sentence as would be recognized by constituent analysis without this restriction. Since, all other things being equal, this will enormously complicate the description,[62] it is necessary to ask for the motivation for the imposition of this constraint on linguistic description. The answer to this lies, I think, in semantic interpretation.

We consider an example discussed by Pike:[63]

> 'Both tagmemic grammar and transform grammar insist that an attempt to describe grammatical structure in terms of morpheme classes alone — even successively inclusive classes of classes — is insufficient. Tagmemics, however, does it by insisting on the necessity of both emic slot and emic class as being relevant to grammatical structure. In this latter we state that an analysis of sentence structures solely in terms of sequences of word classes such as NVN is rejected, and one such as S^N P^V O^N with both slot and filler explicit or implicit .. is essential.'

In the notation used by Elson and Pickett, which we are following, the latter formula would be S:N P:V O:N. Why does Pike feel that this formulation is superior to the NVN formula? The reason is, I think, the assumption that an attempt to describe the semantic features of sentences requires in some sense a semantic interpretation of the nonterminal constituents, that is, of elements like 'Verb', 'Subject', etc., as well as of the terminal constituents or morphemes.

Thus in English if we take John likes Bill, Bill likes John, and all similar sentences which are NVN, we would presumably find that the class of items which occurs in the postverbal position is the same as that in the preverbal position. Yet it would appear that the fomrer is in some sense associated with the meaning of 'subject', the latter

with that of 'object'.[64] But if there is only the representation NVN, there appears to be no way to indicate this fact. Pike thus recognizes a higher order set of distinct constituents, S and O, which can be assigned different semantic interpretations. There are then two different tagmemes, S and O, which have the same filler class N.

Notice, however, that even if one accepts the assumption that a direct semantic interpretation of at least some nonterminal constituents must be provided, it does not follow that Condition (5) must be accepted. For this would be to assume that the only way to distinguish two constituents which are identical in lower order membership, say preverbal N and postverbal N in the example above, is to recognize different higher order categories above them. But in fact a PSG provides an entirely determinate way to do this in another manner by taking advantage of their different positions. There is no reason why the semantic description cannot say that preverbal N is interpretable as 'subject', postverbal N as 'object', etc. A PSG provides a perfectly mechanical enumeration not only of constituents and their labels, but also of their order relative to other constituents.[65] To quote Chomsky:[66]

> 'In terms of such Phrase markers, we can define grammatical relations as certain subconfigurations. Thus the Subject-Predicate relation might be identified as the subconfiguration (Sentence: NP, VP) in which case it would hold between 'John' and 'is easy (eager) to please'.'

It would appear that a failure to study the formal properties of the SD enumerated by PSG in their tagmemic form, a failure made almost inevitable by the lack of precise formulation of these formulas and the way they assign P-markers, has led to the erroneous conclusion that Condition (5) is a necessary result of requiring an interpretation of nonterminal constituents. Notice that since Condition (5) greatly complicates descriptions, even if the need for semantic interpretation of nonterminals is accepted, it follows that the method which utilizes order is to be preferred since it affords a simpler description. Furthermore, the method which takes advantage of order provides an account of such notions as 'subject' which is <u>relational</u> as opposed to Pike's description in which such a notion is simply characterized as a constituent of a certain sort. But it would appear that as opposed to, say, 'Noun Phrase', the notion of 'subject' is inherently a relational one. Thus it makes sense to say that 'X stands in the subject relation to Y', but hardly that 'X stands in the 'Noun Phrase' relation to Y'. The tagmemic characterization

thus seems to miss the relational aspect of grammatical features like 'subject', 'object', 'predicate', etc. and confuses these with constituents.

Tagmemic formulas usually contain a device for indicating the obligatory or optional character of an element in an expansion. A plus sign is used to indicate an obligatory element, a combination of plus and minus signs an optional element. Thus there are formulas like:

$$A = +B{:}b + C{:}c \pm D{:}d$$

which is to be interpreted as an abbreviation for the set of formulas:

$$A = B{:}b \quad C{:}c \quad \text{and} \quad A = B{:}b \quad C{:}c \quad D{:}d.$$

The plus and minus signs are thus equivalent to Chomsky's use of parentheses to compress similar rules, which we have adopted above. In these terms the compressed tagmemic formula could be given as:

$$R31 \quad A \rightarrow B \quad C \quad (D)$$

The use of plus and minus signs thus in no way affects the PSG interpretation of the kinds of formulas to which it is added.

It can be concluded that insofar as tagmemics involves formulas like A = B:b C:c ... N:n it is readily interpretable as a variety of the theory of PSG, furthermore, as a variety of CF-PSG. The imposition of Condition (5) in no way proscribes the PSG interpretation and does not appear to be required by the motivations which have led to it.[67] These are in fact dubious.

An interesting question is whether or not tagmemics countenances contexts in rules. In view of the fact that there are no explicit statements as to the abstract character of the formulas, this must be determined by an investigation of actual formulas. A CS interpretation will be rqeuired if there are pairs of formulas like:

$$A = B{:}b \quad C{:}c$$

$$D = B{:}e \quad F{:}f$$

that is, whenever the same tagmeme has two or more distinct filler classes whose occurrence is dependent on other tagmemes with which it cooccurrs. There does not appear to have been any recognition on the part of tagmemic writers that pairs of formulas like those above involve a fundamentally different form of rule than sets of

formulas in which pairs like B:b C:c and B:e F:f do not occur. It seems clear that, in general, tagmemic formulas are thought of in such a way that a CF interpretation holds.[68] However, I have found at least one case[69] in which pairs like the above are found and thus where a CS interpretation is required. Although there is no space here to support this, I think that a full recognition of the difference between CF and CS-PSG rules, and a recognition of the uses of the latter, would make a great difference and improvement in tagmemic descriptions even within their PSG framework.

The subject of contexts is of sufficient interest to merit a brief theoretical discussion. This can be best given by means of an artificial example. Consider the set of strings

$$a_1X \; a_1BCc_1$$
$$a_2Y \; a_2BEc_2$$
$$a_3Z \; a_3Fc_3$$
$$\ldots\ldots\ldots$$
$$a_nW \; a_nBDc_n$$

Suppose that n is finite and that it is known on independent grounds that those elements before the space should have the structure of nouns, that those elements after the space are as a whole verbs, that a_1-a_n are person markers, and that c_1-c_n are aspect or tense markers of some sort. Clearly, then, a_1-a_n and c_1-c_n should have common representations, say Person and Aspect respectively. Suppose, furthermore, that BC, BE, F, BD, etc. are known to be verb bases and that X, Y, Z, W, etc. are known to be noun stems. All information of this sort can be given by rules of the form:

R32 S → Noun Verb
R33 Noun → Noun Prefix Stem
R34 Noun Prefix → a_1, a_2, a_3, ... , a_n
R35 Verb → Person Base Aspect

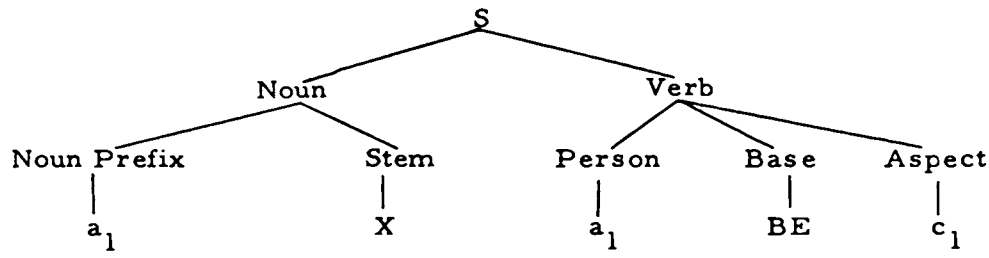
R36 Aspect →
- c_1 in a_1 Stem Person Base _____
- c_2 in a_2 Stem Person Base _____
- .....
- c_n in a_n Stem Person Base _____

R37 Person →
- a_1 in a_1 Stem _____
- a_2 in a_2 Stem _____
- .....
- a_n in a_n Stem _____

R38 Stem → X, Y, Z, . . . , W
R39 Base → BC, BE, F, BD, etc.

These rules will then produce P-markers like:

Diagram 11



R36-7 are CS rules. The set of terminal strings enumerated can, however, be achieved by a set of exclusively CF-PSG rules, namely:

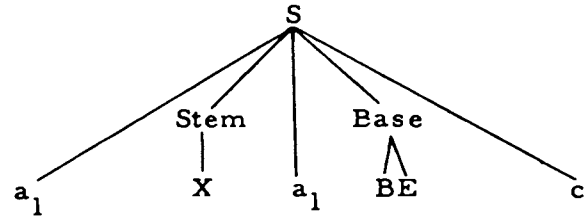
R40 $S \rightarrow \left\{ \begin{array}{l} a_1 \text{ Stem } a_1 \text{ Base } c_1 \\ a_2 \text{ Stem } a_2 \text{ Base } c_2 \\ \ldots\ldots\ldots\ldots\ldots \\ a_n \text{ Stem } a_n \text{ Base } c_n \end{array} \right\}$

R41 same as R38

R42 same as R39

Note, however, that the P-markers produced by these rules are the incorrect:

Diagram 12

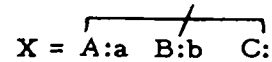


In short, it can be seen that in certain cases it is possible to replace CS rules with CF ones but only at the cost of radically reducing the amount of phrase structure imposed. This is done by eliminating all those intermediate constituents between the set of elements which are context restricted and the lowest node which dominates a string which contains these restricted items. In any real language case the general replacement of CS rules by CF ones will greatly complicate the description because it will force replacement of single items like Verb by whole sets of strings which they dominate. Since constituents like Verb will occur many times in a full grammar, this represents a great loss. The artificial example just given makes the CF description seem simpler only because it is so restricted.

It is important to ask whether any set of CS rules meeting Condition (3) can be replaced by a set of CF rules with the same weak generative power, that is, without changing the set of enumerated terminal strings. It has been proven by Parikh that this is not the case. There are CS-PSG which cannot be reformulated as CF-PSG.[70] Thus, use of contexts, even in conjunction with Condition (3), which is known to limit generative power does add to weak generative power. Parikh's proof consisted in providing an artificial language which was beyond CF-PSG description but not beyond that of CS-PSG. Unfortunately, this language has no linguistic interpretation and consists of such an unnatural set of strings that it does not help very much in trying to determine what sorts of CS rules can be reformulated as CF ones and what kinds cannot. A characterization of CS rules in this regard might be one way of approaching a determination of the weak generative power of CS-PSG.

We see then that there are interesting theoretical issues bound up with the question of contexts in PSG rules and that tagmemics has failed to make clear just where it stands on these. The more general formulations of this view as well as certain particular analyses given in terms of it strongly suggest a CF view.

There is a device used by some tagmemic writers to indicate that one of two elements but not both may occur in the same tagmeme.[71] This is of the form:



Such a notation indicates that the element on the left of the formula consists of either B:b, A:a B:b, or B:b C:c (but not A:a B:b C:c). If there is a plus sign above the line, one of the two possibilities connected by the line is required. In most cases C:c is identical with

A:a but this is not necessary. This notation is then naturally interpretable as a way of abbreviating two tagmemic formulas, namely:

$$X = A{:}a\ B{:}b$$

$$X = B{:}b\ C{:}c$$

That is, this device of a raised line is roughly a suggestion on the order of the use of parentheses and brackets in PSG rules by Chomsky. It in no way alters the PSG characterization of tagmemics. However, I think that in an interesting way this device together with an associated habit has served to prevent followers of tagmemics from discovering one of its inadequacies.

Consider instances of the above device where A does equal C.[72] The chief use of this will be to handle cases of free constituent order. Notice that in a transformational grammar such cases will be handled by generating one underlying base order in the constituent structure rules and then deriving the others by later optional permutation transformations. There are strong motivations for this description. Foremost is the fact that in such constructions of the form ABC, BAC, ACB, etc., the cooccurrence restrictions between the elements are identical for all orders. But if all possible sequences are generated directly by the PSG rules, as in the tagmemic system, these selections must be stated several times. Thus, use of the line device, in conjunction with a strong tendency not to state coocurrence restrictions by means of explicit rules, has obscured for tagmemic writers the fact that PSG description of freely ordered constructions involves multiple statement of identical selections and thus pointless complication of the grammar. Transformations avoid this by applying to P-markers which have all coocurrence restrictions built in. These restrictions are carried over under permutation and need be stated only once for the base order.

It should be noted that at least once Pike has suggested a type of rewrite description which is definitely beyond PSG interpretation. I refer to his suggestion[73] that in certain cases one morpheme be taken to represent two tagmemes. This amounts to deletion. If, however, deletions are allowed in rewrite systems, P-markers cannot be uniformly assigned by the only method known, as was noted earlier in discussing Lamb's system. The motivation for this move on Pike's part is only to avoid having morphemes which are always phonologically null. I find this no justification since it amounts to a restriction on morphophonemic rules that they may be of the form $A\ B \rightarrow I\ B$ but not $A \rightarrow I$. That is, deletions are only allowed in some nontotal set of environments. This restriction will complicate many

morphophonemic descriptions while adding nothing to the validity of grammars.

It was noted earlier that there were thirty-eight tagmemic formulas in Elson and Pickett's introduction and that thirty-five of them had the form A = B:b C:c ... N:n. The other three are concerned with agreement. Since this topic is of general interest, we turn to it now before concluding our discussion of tagmemics.

VIII. Agreement, Variables, and PSG Rules

The phenomenon of agreement or concord is well-known to linguists. The present discussion will be framed about the case of Spanish article-noun-adjective agreement since this is simple and straightforward and because Elson and Pickett attempt to deal with this in tagmemic terms:

One finds:
la alumna 'the pupil (fem)'
el alumno 'the pupil (masc)'
la mujer 'the woman'
el hombre 'the man'
las alumnas 'the pupils (fem)'
los alumnos 'the pupils (masc)'
las mujeres 'the women'
los hombres 'the men'
la alumna buena 'the good pupil (fem)'
las alumnas buenas 'the good pupils (fem)'
el hombre bueno 'the good man'
los hombres buenos 'the good men'
una alumna 'a pupil (fem)'
un hombre bueno 'a good man'
etc.

But no other combinations of these elements with stems constant. Thus the possible morphemic strings in this construction are:

1 F Noun Stem Fem F (Adjective Stem F)
1 F plural Noun Stem Fem F plural (Adjective Stem F plural)

the same set of strings with F everywhere replaced by M and Noun Stem Fem replaced by Noun Stem Masc, and all four of these strings with 1 replaced by un.

Elson and Pickett characterize the situation in an incomplete way and hesitate on the recognition of the gender morpheme in the

noun. They give the formula:

$$N = \pm\ L{:}ar + H{:}n \pm M{:}a_1/a_2$$

and they say:[74]

> 'that is, one kind of noun phrase consists of an optional limiting slot filled by article, an obligatory head slot filled by a noun, and an optional modifying slot filled by an adjective type 1 or adjective type 2, ...'

Adjective of type 2 refers to a putatively irregular group of forms and need not concern us further. They then expand a_1:

$$a_1 = +nu{:}as + g{:}120 \pm pl{:}110$$

Thus for them an adjective of type 1 consists of a nuclear slot filled by an adjective stem, a gender slot filled by an affix of class 120, and an optional plural slot filled by a 110 affix.

They then say that the article and the adjective must agree with the noun in gender and number. Later they claim that:[75]

> 'Note that the agreement feature is shown by the lines in the formula and the detail in the numbered statement below the formula.'

The numbered statement is the one paraphrased at the beginning of this paragraph. What I wish to ask is how in fact the agreement is shown. In other words, what formal interpretation can be given to the use of the connecting lines in the tagmemic formula for agreement given by Elson and Pickett? If the information about cooccurring strings is to be given precisely, then there are two possible ways of interpreting the lines. On the one hand, they may simply be a way of abbreviating a set of PSG rules, that is, a whole group of tagmemic formulas of the type considered earlier. This interpretation can be illustrated as follows. Since I think the tagmemic formulation given by Elson and Pickett introduces a number of unnecessary complexities (because of Condition (5)), I shall give the following ordered rules in my own terms and not attempt to make them a direct translation of their formula. The relations should be obvious.

R43 NP → Article Noun (Adjective)
R44 Noun → Noun Stem Gender 1 (plural)
R45 Article → { 1 / un } Affix 1

R46 Affix 1 → { Gender 2 plural in __ Noun Stem Gender 1 plural / Gender 2 }

R47 Adjective → Adjective Stem Affix 2

R48 Noun Stem → Noun Stem Fem, Noun Stem Masc

R49 Gender 1 → { M in Noun Stem Masc ___ / F }

R50 Affix 2 → { Gender 3 plural in plural Adjective Stem ___ / Gender 3 }

R51 Gender 2 → { M in ___ (plural) Noun Stem Masc / F }

R52 Gender 3 → { M in M (plural) Adjective Stem ___ / F }

R53 Noun Stem Masc → alumn, hombre, etc.

R43 Noun Stem Fem → alumn, mujer, etc.

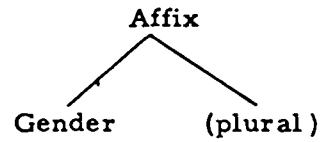
R55 Adjective Stem → buen, etc.

This little grammar provides explicitly for the correct strings of morphemes in the type of Spanish noun phrase under consideration. Notice that from the point of view of this description, which is, I believe, the simplest possible PSG description that provides even the rudiments of a correct constituent analysis, the tagmemic recognition of L, H, and M constituents is entirely redundant. Similarly, the recognition of a plural constituent distinct from a plural morpheme is formally absurd and is not even required by the original semantic motivations which led to Condition (5). Indeed, since a description of Spanish containing a plural constituent follows from Condition (5) (Because rules like R30 above are excluded), this may be taken as a reductio of this condition. Adding all the tagmemically required elements to the above description would greatly complicate the rules, as can be easily determined by the reader.

It is thus clear that the intent of the Elson and Pickett description of agreement can be given by CS-PSG rules (Since tagmemics usually countenances only CF-PSG rules, this is already an important departure.). But obviously there is something very wrong with the PSG description. First, it is forced to recognize three different gender constituents and two different affix constituents in order to expand one of each type first and then the others in terms of the subsequent development of the first. Note, however, that distinct labels make the truth claim of distinct structure (and it might be reasonable to impose the condition that two distinct constituents A and B must dominate at least partially distinct sets of terminal strings.). Therefore, if the identity in structure of all affix

sequences is to be represented correctly, a single set of affix constituents for article, adjective, and noun alike is required, of the form:

Diagram 13



Secondly, the above description is enormously complex. This claim obviously only makes sense if there is a simpler description. But indeed there is, namely, that given by Elson and Pickett themselves in English to the effect that 'the article and the adjective must agree with the noun in gender and number'. As we show in a moment, this quite traditional form of statement can be formalized as a transformation. In PSG terms an essentially separate statement is needed for each agreement, that is, for each pair of morpheme strings which must be the same. The fact that the agreeing strings are instances of some higher order category is a fact which cannot be taken into account. Although the actual agreements are between individual morpheme sequences, any general statement of the agreement must be given in terms of higher order categories. This is only possible in a system in which the categories or constituents can occur as variables. PSG in any of its versions, tagmemics included, is not such a system. However, in TG rules may be stated in terms of general categories like Noun, Verb, Affix, etc., but apply to particular individual sequences of morphemes by virtue of the fact that transformations are defined on analyses of P-markers which act as variables over individual P-markers whose last lines are particular strings of morphemes.

Consider now a TG description of these Spanish noun phrases.[76] Instead of R43-55 the following set of PSG rules is required:

R56 NP → Article Noun (Adjective)
R57 Noun → Noun Stem Affix
R58 Affix → Gender (plural)
R59 same as R48
R60 same as R49 with Gender replacing Gender 1
R61 Article → l, un
R62-4 same as R53-5 respectively

These rules provide no gender or number morphemes for either adjective or article. The grammar must thus contain the following agreement transformation:

$T_{agreement}$
Article, Noun Stem, Affix, (Adjective)
1 2 3 4

1...4 ⇒ Article + Affix, Noun Stem, Affix, (Adjective + Affix)

This transformation adjoins the Affix of a Noun to the Article constituent and to the Adjective constituent if one is present. Notice that although this rule is stated in terms of such higher order categories as Affix, Noun, etc., it will produce equivalences of individual morpheme sequences because these categories serve as variables. Thus $T_{agreement}$ would operate on the P-markers in Diagrams 14 and 16 to produce the outputs in Diagrams 15 and 17.

Diagram 14

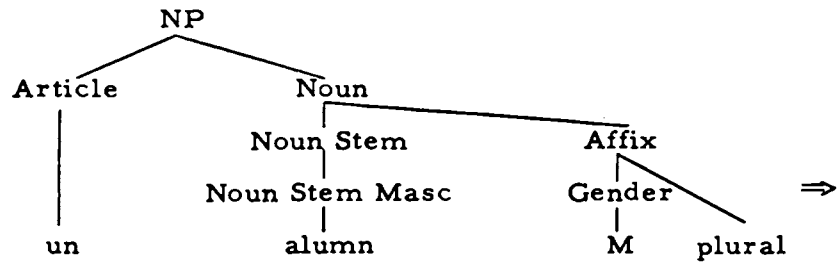
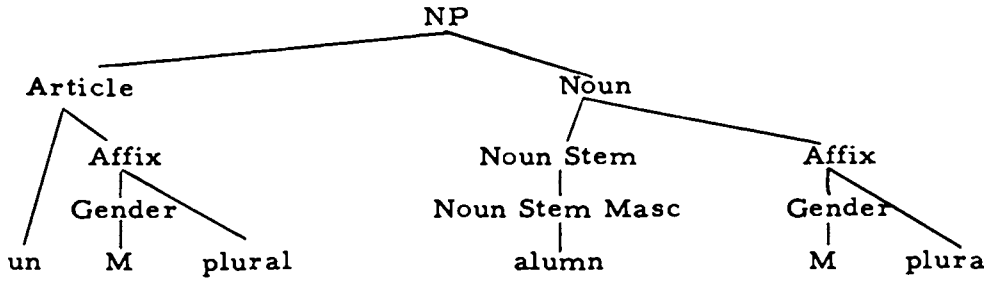


Diagram 15



unos alumnos

Diagram 16

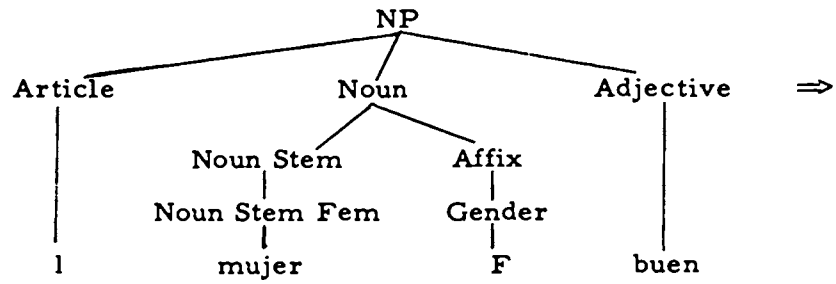
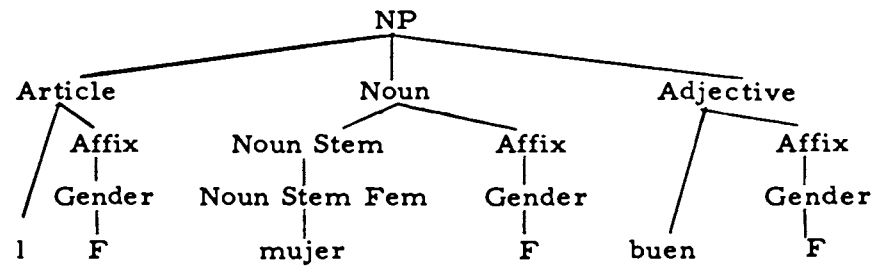


Diagram 17



la mujer buena

It does not require any subtle measure of grammatical simplicity to determine the vast extent to which the description of R56-64 and $T_{agreement}$ is an improvement over R43-55. The transformational description embodies all the generalization possible, accounting for the varied matchings by a single statement. An exclusively PSG description requires at best statements whose complexity is a direct function of the number and length of distinct sequences which must match. This fact produces a great difference in simplicity between the exclusively PSG description and the TG description, even in the above Spanish example where there are really only four different agreeing sequences. As this number rises, the differences in complexity will mount accordingly. For note that $T_{agreement}$ would hold regardless if Affix dominated ten or a hundred distinct sequences of any degree of internal complexity. In other languages besides Spanish,[77] many more necessarily matching sequences can be found, often of significant internal complexity. In Mohawk[78]

there are twenty distinct pronominal sequences which most agree with the verb both as subject and as object and which also must agree with a possessed noun. Many of these contain two, and many three internal elements. Similarly, the well-known concord facts in the Bantu languages are of much greater complexity than those illustrated for Spanish, etc. The advantages in terms of simplicity of TG over PSG descriptions of concord can thus hardly be exaggerated.

It should also be pointed out that the TG description of agreement provides a superior SD since it is not forced to recognize ad hoc constituents like Affix 1, Affix 2, Gender 2, etc. A single set of Affix constituents suffices as is intuitively correct.

It seems clear that tagmemic writers have not considered the formal interpretation of the agreement lines in formulas like those given by Elson and Pickett.[79] Within the general frame of reference of tagmemics, however, the PSG interpretation seems justified. This requires not only contexts but also a very strict ordering of rules if the agreement is to be stated rightly. Neither of these features of grammars have, to my knowledge, ever been discussed by the formulators of tagmemics.

It was noted earlier that there were two possible interpretations for the lines in the tagmemic formulas for agreement. The second is that they are to indicate some sort of variables. Thus in one sense they might be interpreted as unformalized agreement transformations. This interpretation is, to say the least, not very plausible. However, there are, of course, other ways of introducing variables into linguistic description besides the present conception of TG.

One of these which would provide a better than PSG description for agreement would be the assignment of subscript indices to constituents. Thus each constituent which is concerned with agreement could be assigned a subscript which varies over the possible different agreeing sequences. Thus for Spanish there might be strings like:

$$\text{Article}_{G,\ N} \qquad \text{Noun}_{G,\ N}$$

where G varies over masculine and feminine and N over plural or null. There would then be a condition that in any given string the value of all subscripts of the same type must be the same. As noted earlier in discussing Lamb's system, an idea like this seems to be involved in his brief remarks about agreement. A system of

PSG rules supplemented by subscripts was actually used by Chomsky[80] to describe agreement in Hebrew.

When compared with the TG description the subscript approach to agreement appears to have few virtues, although it is certainly better than PSG conception. The fundamental inadequacy of the subscript idea is that it assumes that variables are needed only in the description of agreement. That is, a model of language which contains only PSG rules supplemented by subscripts for the description of agreement treats the latter as an isolated and special feature of language. In the TG it can be shown that the devices needed for the description of agreement are needed for many other aspects of syntax. Thus the TG provides a general account of grammar from which the description of agreement by variables follows. In the subscript system the use of variables must be introduced ad hoc just for this purpose. The subscript conception thus provides much less of an explanation of agreement and the general theory of grammar which uses subscripts would be simpler if this phenomenon did not exist. On the other hand, the theory of TG would be unaffected by the absence of agreement, and it is thus able to show that this aspect of grammar is a special case of deeper regularities in language, namely, the existence of adjunction transformations. Secondly, it might be argued that the SD provided by a subscript description of agreement is inferior since it provides no reason why some morphemic material should have a subscript origin while most derives from nonsubscript constituents.[81]

IX. Tagememics: Conclusions

I have tried to show that the devices supplied by tagmemic writings for the morpho-syntactic description of natural languages are, with one exception, clearly interpretable as PSG rules. The one exception concerns the description of agreement. But even here the PSG interpretation seems most natural. Those features of tagmemics which are considered most distinctive by its developers are its insistence on labelling, its general recognition of branchings of degree greater than two, and its use of 'slot-class correlations'. Far from placing the system beyond PSG interpretation, however, these simply show that it is a variant meeting Condition (5). This condition greatly complicates descriptions and is not required by the motivations which appear to have led to it.

Thus tagmemics is a subtype of the theory of PSG. But it is evidently a formulation of this model of an extremely unclear sort. Questions of structure assignment, ordering of rules, optional versus obligatory character of rules, possibility of permutations, statement of selections, use of contexts, etc. do not appear to have even been raised to say nothing of being answered. In view of this, Pike's claim that:[82]

> '. . . so it would seem possible that if tagmemics and transform grammar are both developed far enough that they could come to a complete overlap.'

cannot be accepted.

The above follows, Pike claims, because TG is an item and process type of description while tagmemics is of the item and arrangement variety and these are equivalent. But this is a misrepresentation of TG. As mentioned in the discussion of Hockett's 'Two Model' article above, from which this contrast comes, IP and IA are indeed apparently equivalent, but both are equivalent to PSG. It is true that tagmemics is a version of the latter. But neither in the weak sense nor in the strong sense is TG equivalent to PSG. PSG provide SD in the form of a single P-marker per sentence. TG provide an SD in the form of a set of P-markers as well as a complex structure of transformations. The present version of TG is certainly not weakly equivalent to any version of the theory of PSG either, although most questions about the generative power of TG remain open. In short, tagmemics is an unformalized version of PSG necessarily containing any flaws possessed by the latter as well as those following from Condition (5) and a lack of formalization or precision. TG, regardless of its truth value relative to PSG about which we say more later, is a radically different sort of theory. The claims made for tagmemics vis-a-vis its relation to TG are thus incompatible with the actual character of tagmemics as a theory of grammar.

X. Hockett's Constructional Grammar

Hockett's article 'Grammar for the Hearer'[83] presents one of the most recent sets of ideas we shall consider. Curiously, these are both the most inadequately formulated as well as those for which the strongest claims are made. The ideas of this article are,

it should be noted, substantially those of his somewhat less recent book.[84]

Hockett proposes what he calls 'Constructional Grammar' (henceforth CG) which he offers as a valid model of grammatical description distinct from both PSG and TG. According to Hockett, CG makes use of four types of element: (1) words; (2) the form classes to which words belong; (3) hierarchical organization or IC (immediate constituent structure); and (4) construction types. The latter notion is defined as 'a set of constructions that have some formal property in common'.[85] Note that the definition is empty since any set of constructions will share some formal property. For example, the constructions represented by in bed (Preposition Noun) and eats quickly (Verb Adverb) share the formal property of having two constituents so that they form a construction type. That is, they are members of the very large construction type containing all and only those constructions with two members, also members of the construction type containing constructions whose first element is a single word, members of the construction type containing constructions with less than 6555 constituents, and to indefinitely many other such grammatically meaningless classes.

Since the notion of construction type is empty, Hockett's applications of it to particular cases are arbitrary. Thus he claims that black cat and ran quickly belong to the same construction type because both are built up by attributives. That is, there is in each case a kind of endocentric expansion of a head which can occur alone. This is a plausible intuitive relation between these two forms unilluminated by the notion of construction type, since this fails to account for the presence of a similarity in this case, and its absence in the case of endless numbers of other construction types, such as the absurdities given in the previous paragraph. Sometimes the notion of construction type is applied by Hockett in such a way that there is even more than one plausible grammatical relation which it might cover. Thus we are told that John is here and is John here both have IC's John and is here put together by two different constructions of the same construction type. He does not specify the construction type further, but besides the absurdities there are the two conceivably interesting and quite distinct possibilities of construction types containing constructions which are full sentences or containing constructions which have identical constituents.

It is apparent what Hockett is attempting to do with the notion of construction type. He recognizes that there are significant relations of various degrees of abstractness between sentences and

parts of sentences. Some of these, such as the notion of 'head' or 'endocentric derivation', are abstract enough to be cross linguistics. The notion of construction type is an attempt to precisely characterize some of these relations. The fact that this notion does not begin to accomplish this task should not, of course, be taken to imply that such an aim is unimportant. In fact, just such explanations should be taken as empirical conditions to be placed on particular grammars and linguistic theory.[86]

We have seen that the notion of 'construction type' can play no role in the characterization of grammatical facts. Consider next the interrelated notions of 'form class', 'IC structure', and 'construction'. We find the following remarks[87]

(1) 'A CONSTRUCTION may be described as a way of putting forms together into larger forms. Abstractly, a construction is a relation: a class of ordered n-ads of forms, where n is always finite, usually small, often exactly 2, but never 1.'

(2) 'A form 'built by' construction is necessarily COMPOSITE; that is, an atomic element (a word, in our current frame of reference) does not belong to a construction.'

(3) 'Two forms, simple or composite, belong to the same FORM CLASS if they have exactly identical privileges of occurrence: that is, if each occurs as the ith constituent in any construction in which the other occurs as the ith constituent.'

(4) 'Two composite forms belong to the same construction if, for all relevant i, the ith constituent of each belongs to the same form class as the ith constituent of the other, and if the two composite forms belong to the same form class.'

(5) 'It follows that every member of a construction belongs to the same form class, but not necessarily vice versa.'

(6) 'The number of constructions is finite.

Hockett's definition of construction is quite vague. Clearly not any class of ordered n-ads of forms is a construction. It is thus best to consider that in statement (1) the term 'form' is taken as primitive to mean roughly 'significant form' or 'constituent'. Thus

eats the in He eats the meat is not a 'form'. In other words, Hockett's notion of construction must be thought of as relativized to some characterization of a significant division or enumeration of sequences of minimal elements. Given this limitation on 'construction', what does the independent notion of IC structure refer to? It is clear from Hockett's usage that he interprets this latter notion so that it refers only to an unlabelled bracketing of elements. Thus for Hockett the pairs in (i) and (ii) have the same IC structure,

(i)	(ii)
runs true	big brook
eats well	nice man

namely:

Diagram 18

But obviously he wishes to distinguish such pairs. And so he recognizes a difference in 'construction'. The first would be 'verb-plus-adverb', the second 'adjective-plus-noun'. But this shows that, regardless of statement (1), Hockett's notion of construction really refers to the labelling of the bracketing. Perhaps more common names for these 'constructions' are NP and VP respectively. It is typical of Hockett's usage that he uses compound terms to denote constructions, where the elements of the compound are the names of the lower order constituents.

The claim that Hockett's usage of 'construction' refers to labelling of the bracketing is apparently inconsistent with his statement (2) that words do not belong to constructions. In a P-marker words are certainly dominated by labelled constituents. However, since for Hockett words do belong to form classes, and since these do have names (Noun, Verb, etc.), our interpretation can hold. Hockett's statement here simply amounts to the decision, which has no theoretical significance in the present context, to differentiate labelling above the word level from that below. He would thus, as

is traditional, not speak of the 'Noun construction', restricting the term 'construction' to sequences of words, or phrases in one sense. Thus far then there are certainly no notions in Hockett's formulation beyond those provided by the P-marker and capable of enumeration by PSG.

Statements (3) and (4) are very important. Between them they state a condition on the set of elements which may belong to a single form class or construction. Notice that only one statement would be needed if Hockett did not draw a line between word level and all lower labellings and those above the word level by distinguishing form class from construction. Together statements (3) and (4) require in terms of P-markers that for two elements to be dominated by the same constituent they must share all environments.[88] But this is to insist that the rules which enumerate constituents are CF. That is, if there are rules:

R65 $A \rightarrow B$ in X_Y
R66 $A \rightarrow C$ in U_W

where X and $Y \neq U$ and W respectively, then B and C cannot share all environments. But if B and C do share all environments, then the rules cannot contain significant contexts. In short, Hockett has imposed Condition (4) on his system which is thus, for the most part, a version of CF-PSG.

There are some damaging deficiencies in the type of SD possible in Hockett's conception of CG. These are most clearly illustrated by his approach to ambiguity and some examples he considers. He says:[89]

> 'Consider the two-word composite form yellow clothes. This can occur in sentence-contexts of either of the following varieties:
>
> (18) Washing in strong soap will yellow clothes.
> (19) She likes to wear yellow clothes.
>
> In a CONSTRUCTIONAL GRAMMAR, we say that yellow is the same word in both, that clothes is the same word in both, but that the two words are put together by different constructions. The first construction is one of the type verb-plus-object, the second one of the type attribute-plus-noun head. Along with this, we are able to say

> that two words are the same word if and only if they are phonemically identical (no homophones or zeros), and that no word belongs to more than one form class. Ambiguity is then handled wholly in terms either of constructions (yellow clothes) or of IC organization (old men and women).'

The last line of this quote is particularly interesting in supporting our view that Hockett's notion of construction refers to the labelling, his IC structure to an unlabelled bracketing. For note that the natural and motivated unlabelled bracketing alone will account for the ambiguity of the phrase old men and women (even in isolation), namely, (old(men and women)) versus ((old men) and (women)), while obviously yellow clothes cannot be so dealt with without labelling. Hockett wishes to account for the ambiguity in this while maintaining that the same words appear in both. Hence he appeals to a difference in construction and calls the first 'verb-plus-object', the second 'attribute-plus-noun-head'. But the ambiguity in such a sequence is accounted for simply and automatically by the natural labelled bracketing since the two sentences will have P-markers something like:

Diagram 19

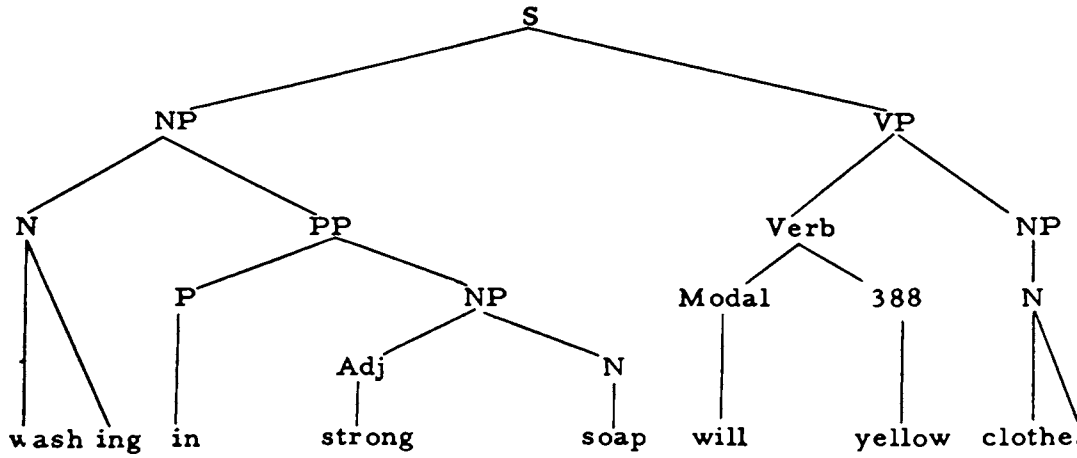
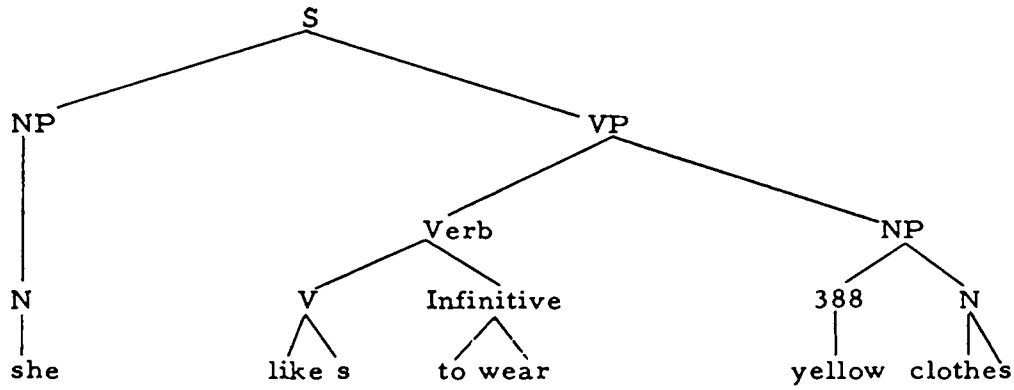


Diagram 20



The arbitrary symbol '388' is used in Diagrams 19 and 20 to indicate the constituent which has as members yellow as adjective and yellow as verb. It is not clear to me that this form class has any other members.[90] The intuitively reasonable description which would be given by most linguists, namely, to replace '388' by 'V' in Diagram 19 and by 'Adjective' in Diagram 20, is impossible for Hockett because he does not allow the same word to appear in two different classes. This whole discussion on Hockett's part is in fact motivated primarily by the desire to illustrate this principle.[91] But note that the result of this assumption is a pointless complication of the grammar on the one hand and incorrect SD on the other. Thus now not only will the grammar need rules like:

R67 Verb → Modal V
R68 NP → Adjective Noun
it will have to add:
R69 Verb → Modal 388
R70 NP → 388 Noun

and so on in countless other cases.

Furthermore, consider the sentences (20) washing in strong soap will destroy clothes, and (21) she likes to wear nice clothes. These can, of course, be represented by P-markers just like those of sentences (18) and (19), replacing yellow in Diagram 19 by destroy, and yellow in Diagram 20 by nice. That is, sentences (18) and (20) and

(19) and (21) certainly have identical structures at all but the lowest level. But note that this analysis implies letting yellow be in two different classes, V and Adjective. This is just what Hockett wishes to avoid. This has, however, the strongly counterintuitive result that (20) does not have the same structure as (18). That is, in (18) yellow clothes is not a member of the same construction as destroy clothes in (20) since, as Hockett's statement (4) tells us, two composite forms belong to the same construction only if each of their members in order belongs to the same form class. But yellow does not belong, for Hockett, to the same form class as destroy, since they certainly do not have identical privileges of occurrence. That is, there is no *she likes destroy clothes. The inadequacy of these descriptions is partially concealed by Hockett's use of the terms 'verb-plus-object' and 'attribute-plus-noun-head' for yellow clothes in its two different functions. For if these terms are used for this phrase, quite different constructions must be recognized in destroy clothes and nice clothes. The proper terms for Hockett would be something like '388-plus-object' and '388-plus-noun-head'.

In short, Hockett's requirements on form classes are such that nothing like the significant groupings wanted for language description can emerge. It is difficult to exaggerate how little the classes which would result resemble sensible grammatical groupings. This lack of correspondence between correct results and the output possible with Hockett's restraints is further disguised by the fact that Hockett contradicts himself. Thus his use of the notion 'subclass' permits a not unreasonable treatment of some examples.[92] But note that, within the complete substitutibility framework of nonoverlapping classes, it is in principle impossible to have any subclasses. For of course different subclasses do not share all environments. Thus the inadequacy of the theoretical constraints on constituents suggested by Hockett is somewhat concealed by the fact that they are ignored in cases where they would lead to unacceptable results. In terms of the theoretical framework one cannot speak of subclasses of plural and singular nouns. There is in this approach no form class of nouns but at best two different form classes of singular and plural nouns. Actually, there could not obviously even be these but instead an unknown number of distinct classes differentiated by every possible distributional property of nouns in English.

The fact that this glaring inadequacy of one hundred percent substitutibility is missed by Hockett is startling in view of the fact that it has been noted again and again over the years by various linguists of radically different persuasions.[93]

It is Hockett's view that the principle that no word belongs to two or more form classes is, along with the possibility of discontinuities, what differentiates CG from what he calls 'phrase grammar' which is a reference to Chomsky's discussions of PSG. He says:[94]

> '(Or, equally well, a single word yellow is assigned to two different form classes, and is said to be a member of one in the one context, of the other in the other.) This alternative deals with ambiguity in terms of homophony or of form classes (yellow clothes) and of IC organization (old men and women). It eliminates constructions by reducing them all to a single 'construction' often called COLLOCATION. With certain further (optional) adjustments -for example, as insistence that forms in collocation must always be adjacent, never separated - this procedure yields what has been called a PHRASE GRAMMAR.'

But this is confusion. The result of eliminating Hockett's insistence that every word belongs to one class is not an elimination of 'constructions,' that is, of labelling. It only amounts to permitting the labelling to be reasonable for a large number of otherwise incorrectly handled cases and eliminates many pointless and ad hoc classes which complicate the description enormously. Note that, given previous remarks, the last sentence in the above quote amounts to an admission that CG is simply PSG with the addition, in some imprecise way, of discontinuities. Our interpretation of Hockett's views is then, in effect, supported by a direct statement.

Hockett then goes on to say:[95]

> 'Phrase grammar (as commonly understood) has a number of known weaknesses, the recognition of which has been the main motive behind the development of transform grammars. Constructional grammar does not share these weaknesses. (emphasis mine: PMP) In the first place, it is possible (as we shall see presently) for constructional grammar to handle quite simply those phenomena that any reasonably manipulable phrase grammar best leaves for a transformational overlay. In the second place, in constructional grammar it is possible to accept the intuitively desirable notion that some words, such as and or or, are not constituents at all but rather MARKERS OF CONSTRUCTIONS ... Finally, and specifically with reference to grammar for the hearer, the allocation of ambiguity to IC organization and to constructions and the elimination of homophony at the word level, seem to yield a kind of realism that a phrase grammar, or a phrase-and-transform grammar, can deal with only in a more indirect and complex way.'

In view of what has been said, it is clear that these claims are without foundation. CG is simply an unclear formulation of CF-PSG with the addition of an unformalized use of discontinuities. Not only must CG share any flaws possessed by the full class of PSG, it has several others of its own due to its insistence on lack of contexts (it is shown later the extent to which this is a flaw) and the inability of the same element to be dominated by different constituents. As we have seen, the kind of 'realism' provided by CG is that in which will yellow clothes has a different structure than will destroy clothes, in which there are no classes of nouns, verbs, adjectives, etc., in which hear and here must represent the same constituent, etc. Furthermore, both CG and PSG handle ambiguity in the same way, namely, by labelled bracketings.[96] The only difference is that within CG the labellings are to a large extent necessarily ad hoc and intuitively incorrect.

What then of the notion of 'marker'? First, clearly if this notion is characterizable by CG it is also by PSG. But Hockett has given no such characterization. It is not enough to simply give the name 'marker' ad hoc to a few words in various sentences. What must be shown is just how the formal devices of the theory assign this status (whatever it is exactly) to and, or, etc. but not to John, runs, quick, etc. in the infinite set of sentences in which they occur. This Hockett has not begun to do and within the limits he imposes no way of doing so appears possible.[97] Thus independently of the virtue of the notion 'marker', Hockett has not provided any account of it and it cannot count in favor of his view and against TG.

Later in this paper, Hockett tries to show successively that TG are 'reducible' or 'convertible' to CG, CG to PSG, and thus PSG to TG. Presumably, although he does not say so, the notion of 'reducibility' is meant to imply weak equivalence, that is, identity of the sets of terminal strings which may be enumerated.[98] First, note that in linguistics weak equivalence as such is of very little importance. To provide support for a grammatical model it must be shown not only that it can enumerate the correct set of terminal strings, but also that it can automatically associate a correct SD with each. These problems are not even raised by Hockett. In this regard, it is important to note that the motivations which originally led to the introduction of TG, as in Chomsky's Syntactic Structures, had nothing to do with limitations on weak generative power of PSG. They were instead that Chomsky claimed that descriptions in terms of PSG provided inadequate SD and were excessively complex. In an earlier remark we quoted from Hockett he claimed that CG does not share the weaknesses of PSG which led to the introduction of TG (presumably

those which we discuss below in Chapter 7). But it is striking that he nowhere mentions any of the inadequacies of PSG claimed by Chomsky, nor does he even attempt to show how CG avoids them.

But second, the so-called demonstration of the weak equivalence of TG and PSG by way of CG is simply fallacious. The present version of TG, the one presented in the available literature and the one with which Hockett is presumably familiar, is powerful enough to enumerate any recursively enumerable set, whereas PSG, to which Hockett claims CG is equivalent,[99] enumerates only recursive sets and not all of these.[100] The actual generative power of CG is, of course, indeterminable since Hockett has provided no account of the types of rules it countenances. From what has been determined about the ideas underlying this conception, however, it is clear that the rules which could be allowed are at best of two types, CF-PSG rules and some as yet unformulated type to provide for discontinuities. Therefore, any difference in weak generative power, and indeed any difference at all between CG and PSG is due entirely to these unformulated kinds of rules. Naturally, then the weak generative power of CG cannot be seriously discussed until these are stated precisely.[101]

XI. String Analysis

In a just published book,[102] Harris presents a system of syntactic description which he refers to as 'string analysis'. Although he distinguishes this from what he calls 'constituent analysis', I wish to argue that string analysis is still another version of PSG.

This kind of sentence description is outlined as follows:[103]

> 'String analysis characterizes the sentences of a language as follows: Each sentence consists of one elementary sentence (its center), plus zero or more elementary adjuncts, i.e. word-sequences of particular structure which are not themselves sentences and which are adjoined immediately to the right or to the left of an elementary sentence or adjunct, or of a stated segment of an elementary sentence or adjunct, or of any of these with adjuncts adjoined to it. An elementary sentence or adjunct is a string of words, the words (or particular sequences of them) being its successive segments.'

Harris first distinguishes 'string analysis' from 'constituent analysis' by observing that in contrast the latter divides a sentence into constituents 'at a lower descriptive level', implying then that string or adjunction description characterizes elements in terms of other features 'on the same level'. He then says the following:[104]

> 'The transition from constituent to string analysis is given, . . . by the observation that most constituents either consist of a single word (of some category, or of any one of several categories which characterizes that constituent) or contain a single word of the characterizing category plus adjunct words or phrases adjoined to it. We can thus consider such a pluri-word constituent in any sentence A as being endocentric, i. e. expanded from its characterizing category by the addition of adjuncts; and this in the sense that we can replace each constituent by its characterizing category alone, and obtain a sentence B which would be related to A as a constituent expansion of A.[105] That is, given a sentence or constituent C whose immediate constituents (i. e. the next-level constituents into which C is decomposable in a regular way) are a C_1-phrase, a C_2-phrase, etc. we find that most C_i-phrases consist of a word of the C_1 category (which characterizes the C_i-phrase) plus zero or more adjuncts of C_i.'

And:[106]

> 'A constituent analysis is replaced by a string analysis when, given a sentence or constituent C whose immediate constituents are endocentric C_1-phrase, C_2-phrase, etc., we define the word sequence $C_1 + C_2 + \ldots + C_n$ as the elementary string of C; and the adjuncts included in the C_i-phrase as adjuncts into the elementary string to the right or left of C_i. An exocentric constituent C_x of C is defined as a string which is itself an elementary segment of the string C, rather than being an expansion of a segment of C'.

Finally, the respective empirical claims of 'constituent analysis' and string analysis are contrasted:[107]

> 'We thus see that the relations formulated in constituent analysis can be included in string analysis, the latter being in various respects more general than the former. String analysis is the stronger of the two, in making the claim that for each class (or for many classes) {C} of constituents there exist elementary members of {C}, and in particular that if the sequence of phrases $C_1P + C_2P + \ldots + C_nP$ is a member of {C}, so is the sequence of word-categories $C_1 + C_2 + \ldots C_n$. In contrast, constituent analysis makes the claim, lacking in string analysis, that the members of each class of sentences (that is, each constituent: PMP) have identical segmentation into constituents, i. e. that the expansion and replacement of sentence segments are encapsulated within a fixed structural segmentation of all sentences of the class.'

It will be remembered from the discussion of Harris' use of raised superscripts in his 'morpheme to utterance' formulations that the notion of endocentricity played a key role. Indeed, we concluded that superscripts were introduced largely for the purpose of characterizing the notion of 'head', which is intimately connected to that of endocentricity. The validity of this conclusion is greatly strenghthened by the great attention paid to such ideas in the new work under discussion and well-illustrated by the quotes just given.

We noted before that Harris' use of raised superscripts involved the claim that all constructions were pseudoendocentric, that is, that the head was always substitutible for the expansion. Harris now abandons that idea and takes 'constituent analysis' in such a way that there can be no 'heads' since all elements of the same structure, that is, in P-marker terms all elements with the same label, must have a fixed division into segments. This means that all occurrences of a given labelled constituent must dominate an identical number of elements. Thus for Harris two subrepresentations in the same description like:

Diagram 21

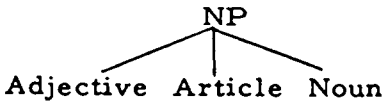
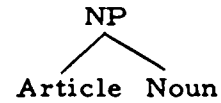


Diagram 22



are impossible in 'constituent analysis' and instead the notion of 'adjunct' is brought in to cover just such cases. In the above diagrams Adjective would be taken as a left adjunct. In short, for Harris 'constituent analysis' now covers just the description of sentence structure up to the point where each constituent is expanded into a head, that is, up to the point where each phrase is characterized by a single word or morpheme.[108] Adjunct or string additions then serve to expand these basic elements by providing some of them with endocentric expansions.

But now in PSG terms it is easy to see that Harris' distinction between 'constituent analysis' and 'adjunct analysis' really amounts to just a differentiation between those elements which occur in parentheses (or, in tagmemic terms, with a plus and a minus sign) from those which do not. Thus if there is a rule like:

R71 $A \rightarrow (B)\ C$

then C is the characterizing category of the phrase type A and B is a left adjunct. In the case where R71 is a low order rule, C will be the characteristic word class of the phrase type A. Thus if a set of PSG rules is provided with conventions for using parentheses, as is necessary on quite other grounds, it is clear that the distinction drawn by Harris is automatically characterized. The so-called 'center' sentence for any sentence is just the string provided by deleting all those elements dominated by any constituent which was parenthesized in its rule of introduction. Thus if nothing else is said, the SD required by Harris is determinable from the set of PSG rules (but not necessarily from the set of P-markers which these produce since these do not reflect the difference between parenthesized and unparenthesized constituents directly).[109]

Harris does not provide any explicit statements about the form of grammatical rule corresponding to string analysis. But Herzberger, in an article attempting to describe the descriptive system behind string analysis developed at the University of Pennsylvania for 'computable syntactic analysis', which is presumably the same system discussed by Harris, makes the following interesting remarks:[110]

> 'The two operations, adjunction and replacement exhaust the recursive aspects of segmental structure.' 'The form of grammatical statement for adjunction is a combinatorial production which develops a chosen letter into a string containing it; that for replacement develops a letter into a string which does not contain it.'

In the terms of the present paper, a 'combinatorial production' is a rewrite rule. Thus Herzberger makes it clear that 'replacement', which would be the form of statement corresponding to Harris' 'constituent analysis', consists simply of PSG rules of the kind which meet Condition (2)c'. 'Adjunction', on the other hand, involves just exactly PSG rules which violate this, namely, those of the form:

R72 $A \rightarrow X\ A\ Y$

Thus if Herzberger is correct, adjunction analysis involves exactly the kind of rules which may lead to nonuniqueness of P-marker assignment. Furthermore, note that rules of the type R72 suffer from the same lack of intermediate symbols which was noted as one reason for the superscripts in Harris' 'morpheme to utterance'approach.

That is, consider adjunction treatment of such a sequence as (Verb, Adverb). If Herzberger is correct, this will have to consist of rules like:

R73 Verb → Verb Adverb

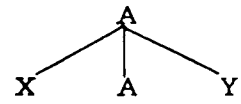
And yet this process would not appear to be recursive. Thus in these terms it will be quite difficult to control recursions properly Notice that in ordinary PSG terms such analyses can be handled by use of intermediate constituents such as VP:

R74 VP → Verb Adverb

But this violates the claim that adjunction involves rules which develop a symbol A into a string Z which contains A.

If Herzberger's formalization is correct, and, as just mentioned, it is quite compatible with Harris' tendency to eschew intermediate constituents, then not only can a PSG characterize the notion of adjunct by means of its rules and a possible complex matching of P-markers, but the notion of 'adjunct' can be read off directly from individual P-markers. That is, given the fact that all adjuncts are introduced by rules of the form R72, adjuncts are simply the elements X and Y in all configurations of the form:

Diagram 23



that is, where there is an element A dominating a string containing A.

It is interesting that Harris' work for a number of years has revolved quite crucially around the interrelated notions of 'head', 'modifier', 'endocentricity', etc. These are important notions and ones which we should certainly require that general linguistic theory characterize correctly. Yet it cannot be held that Harris' work has accomplished this. It would seem that his analysis has not gone beyond a suggestion equivalent to the claim that the notion of 'head' may be characterized in terms of nonparenthesized elements and perhaps to the claim that 'adjuncts' or nonheads may be characterized in terms of P-markers like Diagram 23. Yet the latter claim cannot be accepted simply because the resulting SD are so strongly counter-intuitive. If man is a Noun, then big men certainly is not. This is generally recognized by saying that both are noun phrases which contain the noun man. Secondly, the latter claim cannot be accepted because of the resulting possibility of nonuniqueness of P-marker

assignment. Finally, the argument for a transformational treatment of 'adjuncts' like adjectives is compelling, as Harris was one of the first to note.

The former claim involved in Harris' analysis, that based on the notion of distinguishing optional from obligatory constituents in expansions, can be accepted. Yet, as has been noted by Lees,[111] this is by no means sufficient. Not all parenthesized elements are instances of modifiers nor are all unparenthesized elements in an expansion which contains parenthesized elements necessarily heads. Lees[112] attempted to strengthen an analysis of notions related to 'modifier' by bringing in the idea of 'major constituent', that is, one which dominates an infinite set. It would seem that further refinements are necessary (involving, perhaps, the notion of 'lexical constituent') and that at the moment linguistic theory cannot claim to have completely characterized any of the notions of 'head', 'modifier', 'endocentricity', etc. [113]

It seems likely that any formalization of Harris' conception of string analysis would require PSG rules with contexts. Harris sometimes speaks of certain adjunctions as 'nonrepeatable'. Making such restrictions precise would presumably require contextual conditions in the rules such that the head category meets the condition which is not, however, met after one application of adjunction, that is, after one application of a rule of the form $XAY \rightarrow XUAWY$.

We conclude by emphasizing that the ideas of string analysis remain completely within the domain of PSG characterization. Nothing either in the type or rules suggested or in the types of SD discussed go beyond rewrite rules and P-markers respectively. String analysis is not a new alternative to TG[114] but simply another version of the theory of PSG whose purported inadequacies orginally motivated the development of TG.

CHAPTER 5: DISCONTINUITIES, REWRITE RULES, AND P-MARKERS

We have seen in considering a number of distinct conceptions of grammatical structure that, besides a lack of clarity in formulation, the chief fact which prevents these from being strictly equivalent to PSG is the recognition of discontinuous constituents. The question of discontinuities thus merits a brief discussion of its own.

The treatment of discontinuities in the writings that have been considered provides in microcosm a picture of the general approach of classificatory linguistics to grammatical questions. As we have seen, a large number of linguists have recognized that sentences do contain discontinuous constituents, that is, that in some sense the relation 'is a' holds for nonadjacent elements. Yet as far as I know, none of those linguists who believe transformational rules unnecessary has attempted to state precisely the type of rules which can assign discontinuous constituents. Nor have they for the most part even given a precise account of the type of SD involved in discontinuities.

One of the virtues of TG is that it provides a straightforward formalization of the notion 'discontinuous constituent'. In TG, discontinuities are for the most part produced by the operation of permutation transformations.[115] That is, if in some sentence there is a sequence DAE and D and E are discontinuous constituents of some higher order constituent B, then there is some P-marker for that sentence in which D and E are continuous constituents of B, that is, in which (X) D E (Y) is a B. There are significant advantages to describing discontinuities in this way.

First, the notion of SD in terms of P-markers is retained. That is, linguistic theory need not add some ad hoc kind of SD just to handle discontinuities. Correct SD for discontinuities are provided simply by allowing sentences to have sets of P-markers. This, as will be noted below, is required on many other grounds and can be made an automatic result of the application of transformations. Since there is a mechanical procedure for assigning underlying P-markers by PSG rules, and since there is a fairly simply algorithm for assigning a new derived P-marker when a permutation transformation is applied,[116] TG provides a formalized enumeration of the SD for discontinuities

simply by characterizing permutation transformations. But these are needed for many other language phenomena such as 'free' word order. Thus again, as in the case of agreement, TG shows phenomena which appear quite exceptional from the PSG point of view to be a special case of more general features of language, namely, the existence of permutation transformations and the fact that SD involve minimally sets of P-markers.

A second great advantage of describing discontinuities in the TG way is that a formal basis is provided for intuitively related structures. Thus the relation between John will kill Mary and will John kill Mary is explicated since they have the same underlying P-marker, the latter being formed by the permutation question transformation which produces the discontinuous verbal will . . . kill.

Notice that not all permutations yield discontinuities. These are only produced where at least one of the elements to be permuted is itself not dominated by some higher constituent. Thus a permutation operating on A, D will derive discontinuities in Diagram 24 but not in Diagram 25.

Diagram 24

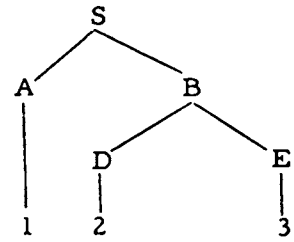
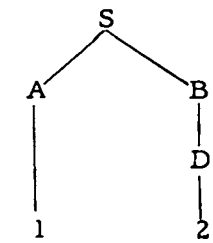


Diagram 25



A third great advantage of the TG description of discontinuities by permutations is that, in cases like the question described above, it accounts automatically for the fact that the selections are identical in both cases because the permutation operates on the underlying P-marker in which these have already been correctly specified. The rule then carries them over to the permuted and hence discontinuous forms.

We have noted that many linguists have recognized the existence of discontinuities. But with the exception of the developers of TG there has been no formulation of a conception of linguistic rules which can automatically assign discontinuous structure. Thus TG appears to be the only theory which accounts for the type of structure

recognized even by those linguists like Hockett who reject TG. Furthermore, the description provided by TG is intuitively natural, preserves the notion of SD in terms of P-markers, is in particular cases formally motivated by the need to avoid duplication of selectional statements, and automatically accounts for intuitive relations between sentences in many cases.

The only attempt I am aware of to describe discontinuities formally in nontransformational terms is that of Yngve.[117] He has attempted to build a generative phrase structure grammar which derives sentences in such a way that discontinuous constituents are produced. He allows rules of the form:

R75 $A \rightarrow B \ldots C$

When these are applied to strings like ADE the result is the sequence BDCE. But Yngve has not presented any algorithm of structure assignment which insures that, given derivations using such rules, it is possible to determine that the 'is a' relation or its equivalent holds between the proper sets of elements and no others. Thus his formalization is far from complete. As noted earlier, the dot notation suggested briefly by Lamb probably would amount to allowing discontinuous rules of the form R75 (as well as those which work in the opposite direction, that is, which when applied to a string EDA will yield EBDC). But Lamb has not even provided the degree of formalization reached by Yngve's approach.

The weak generative power of discontinuous rules of the above types is largely unknown. Recently, Matthews[118] has proved for a larger class of discontinuous type rewrite rules, of which the above are a special case, that their weak generative power is not stronger than that of CS-PSG rules which do not meet Condition (3), that is, which allow permutations.

However, independently of the weak generative power of discontinuous rules and independently of the fact that no one has provided them with an algorithm of structure assignment, it is clear in terms of both SD and simplicity that, with respect to the description of discontinuities, such rules will be far inferior to TG. For like other rewrite rules which do not operate with variables, discontinuous rewrite rules will not provide a method for avoiding duplication of selectional statements nor will they provide a formal basis for the intuitive relations between sentences such as John will go and will John go. Like PSG description, discontinuous rewrite rules can provide no underlying P-marker. Sentences like the above will

simply have distinct derivations as in a PSG and will not be any more closely related than John hit Bill and John hit Harry. Furthermore, of real theoretical import is the fact that description of discontinuities by rewrite rules forces abandonment of the traditional, intuitively correct, easily formalized, and generally agreed on view that the SD of a sentence is, on the level of constituents, definable in terms of the notion of P-marker. TG retains this notion by providing sets of P-markers for each sentence. The use of discontinuous rewrite rules is forced to claim that in cases of discontinuous constituents, but not others, SD must be characterized in some (as yet ungiven way) distinct from P-markers.[119] This pointless ad hoc view of sentence structure is avoided by TG.

To conclude this discussion of discontinuities we should take note of an unfortunate tendency which has crept into some tagmemic writings. This is the idea that a description is better to the extent that it avoids the recognition of discontinuities.[120] This strange assumption follows in a way from the general tagmemic tendency (in the absence of underlying P-markers) to decrease the amount of structure associated with sentences by recognition of multiple branchings.[121] Thus if there is a string ABC in which A and C are discontinuous constituents, recognition of a ternary analysis naturally simply does away with the problem of recognition of discontinuities. But this is only to force the facts of language into an inadequate conceptual mold. Tagmemics, like other versions of PSG, has no natural place for discontinuities. But the existence of such can hardly be doubted, certainly not by anyone who has, for example, ever looked at a German sentence of any degree of complexity.

CHAPTER 6: OTHER SYSTEMS

We have tried to show that eight post war American conceptions of grammatical structure are, in essence, versions of PSG in the sense of Chomsky. It should be noted that there are a number of other systems of linguistic relevance which have been shown to be versions of PSG. In particular, the so-called 'categorical grammars' of Bar-Hillel have been shown to be weakly equivalent to CF-PSG.[122] Chomsky and Schützenberger[123] have proven that the so-called 'pushdown storage' or 'predictive' sentence analysis approach of Oettinger, Sherry, Rhodes and others[124] is weakly equivalent to CF-PSG. That is, any language accepted by a pushdown storage can be enumerated by a CF-PSG and conversely. Recently, Gross[125] has shown that the so-called dependency model approach suggested by Tesnière and elaborated by Hays[126] and others is equivalent to CF-PSG in the weak sense, that is, the set of languages enumerable by the dependency model is just the set enumerable by CF-PSG. Finally, Matthews[127] has demonstrated that any set of rules of the general PSG type, with or without discontinuities, is weakly equivalent to CF-PSG if it is insisted that symbols be expanded in a unique order, either left-to-right or the reverse. Yngve's system[128] falls into this class since he insists on left-to-right expansion.

CHAPTER 7: THE INADEQUACIES OF PSG

In previous chapters we have argued that eight important and characteristic conceptions of descriptive grammar formulated in the United States since the war are, in essence, versions of PSG. It was also shown that a variety of approaches to grammar on the periphery of linguistics, most originating in machine translation work, are versions of the theory of PSG, in fact of the subbranch concerned with CF-PSG. I should like to very strongly emphasize that showing of a grammatical model that it is a version of PSG is to show, in effect, that it is incorrect. PSG are simply incapable of correctly characterizing the sentences of human languages. This conclusion was the basis for the development of TG. I should note again that strong as this claim is, it has not in general been explicitly disputed by linguists. Indeed, such linguists as Pike, Hockett, and Harris, among those we have discussed, have in fact admitted the existence of shortcomings in the theory of PSG.[129] However, certain of these linguists, most notably Hockett and Pike,[130] have maintained in effect that the forms of grammar they propose are not phrase structure and are thus not affected by the inadequacies of PSG. These authors, and many others, thus feel no need to accept the conclusion reached by Chomsky that the inadequacies of PSG require a strengthening of linguistic theory to include transformational rules, because they view their own descriptive devices as legitimate alternatives to TG not subject to the limitations of PSG. However, as shown in Chapter 4, contemporary nontransformational conceptions of grammar are versions of the theory of PSG and thus suffer from whatever flaws inhere in this theory, as well as certain other special ones which were noted in the appropriate place.

The counterevidence which has been brought forward against the theory of PSG is overwhelming. Although there is not the space here to discuss in detail all the weaknesses of PSG, it is important to list those points where PSG fail and to discuss in slightly greater detail a type of counterevidence to this theory which is rather recent and which has not hitherto been available to the general linguistic public.

1. PSG proveably assign the wrong P-markers to coordinate constructions. This was discussed above in our treatment of Wells' ideas.
2. In an enormous number of cases the correct SD of a sentence requires assignment of a set of P-markers. PSG in all cases can assign at best one P-marker. One type of case was discussed in our treatment of discontinuities. Others are worth mention. One of the most common is so-called anaphora in examples such as Bill will go and so will I, John is bigger than I am, etc. PSG can provide no grounds in such cases for explaining why an element or sequence of elements in the first part of these sentences is understood to be repeated in the latter half although not actually present in the utterance. Similarly, PSG can provide no account of the fact that English imperatives like go home, etc. are understood to have a 'you' or second person subject. That is, PSG cannot account for deletions. In a TG, however, it is easy to show that such sentences are derived from underlying P-markers in which the understood elements are actually present. Another type of phenomenon requiring a set of P-markers for adequacy of SD are the so-called substitutions. PSG provide no basis for the fact that forms like he, her, itself, etc. are understood as replacing whole sets of elements. The fact that substitutes have a common distribution with the forms which they are understood to replace shows that to just this extent they cannot be differentiated formally from the replaced elements by sub-class techniques. But in a PSG substitutes can only be considered a subclass of forms entirely parallel to other subclasses which are, however, not understood as substitutes. In particular, English pronouns can only be considered a subclass of nouns like proper nouns etc. Thus the substitute character of pronouns and nonsubstitute character of proper nouns is unexplained. In a TG, however, it is easy to show that pronouns (but not proper nouns) are in fact introduced by substitution transformations from

underlying P-markers in which members of the replaced classes are found. TG thus provide a formal basis for the 'substitute' character of pronouns and similar forms and simultaneous explanation for the absence of this feature in proper nouns, etc. Many other examples of the need for SD to contain sets of P-markers rather than single P-markers can be found.[131]

3. PSG fail to meet reasonable requirements of simplicity. This flaw can hardly be exaggerated. Many easily formulated and quite traditional sorts of rules are excluded, whole sections of the grammar must be repeated several times, etc.[132] One type of simplicity argument was discussed above in our treatment of agreement.

4. PSG fail to account in a non-ad hoc way for the formal structure underlying speaker's knowledge of grammatical relations. Certain aspects of this were discussed earlier in the treatment of tagmemics. For further discussion, cf. pages 111-112 of the Appendix, below.

5. PSG completely fail to account for intuitive relations among sentences such as active and passive, question and declarative, assertion and negation in English, incorporated and nonincorporated, plain possessive and possessive agreement in Mohawk, etc.[133]

6. The notion of sentence type is not reconstructed.[134]

7. In many cases where the correct P-markers can be provided, they are formally unmotivated. That is, the simplest PSG of the sentences would provide the wrong P-markers.[135]

8. PSG cannot handle overlapping selectional restrictions without extraordinary and inelegant complexities and repetitions. These complexities are due to the cross-classifications of categories found in language. For example, the Noun Stem constituent in Mohawk must be divided into classes with respect

> to the properties possessable-not possessable, incorporable-not incorporable, animate-inanimate, etc. Whichever division is made first in a PSG, the second will have to be given twice, the third at least four times, etc. With context-restricted categorizations the situation is even worse.[136]

Flaws 1-7 can be summed up by saying that PSG fail to provide a correct account of SD and fail to provide for reasonably simple description. It has been shown in the literature how each of these flaws is overcome by the use of transformational rules. A few illustrations of this were given in our earlier discussions. Flaws of the 8th type cannot be overcome by transformational rules since they are found in the part of the grammar which must be correctly characterized before transformations apply. At the moment there is, I think, no completely adequate way of dealing with them. It is important to note, however, that this flaw was discovered only through the attempt to utilize explicit PSG rules. It is striking that none of those writers considered above who suggest imprecise versions of PSG have discovered this quite basic inadequacy.

Note that none of 1-8 above have anything to do with weak generative power. Indeed, when Chomsky claimed that PSG rules were inadequate and required supplementation by rules of the transformational type, he was careful to point out that he did not know if PSG were or were not adequate from the point of view of weak generative power.[137] But recent work permits us to add facts from the domain of weak generative power to the counterevidence to the theory of PSG.

It is now known that natural languages cannot even be weakly generated by CF-PSG. This claim was proven by showing that Mohawk, one of the remaining Norther Iroquoian languages, is not a CF-PSG language. That is, it was demonstrated that it is impossible to construct a finite set of CF-PSG rules which will enumerate all and only Mohawk sentences.[138] It was proven by Chomsky[139] that the language $/\gamma\gamma/$ is not a CF-PSG language, where γ is a variable over strings of arbitrary length in a vocabulary of two or more symbols. The fact that Mohawk lies beyond CF-PSG description was shown by demonstrating that Mohawk contains an infinite subset of sentences with the formal properties of $/\gamma\gamma/$. This subset is based on the fact that Mohawk sentences may contain both incorporated noun stems in the verb and external noun objects containing a noun stem. These noun stems must be identical and there is a recursive process for deriving noun stems from verb bases which may include incorporated noun

stems. Thus there is no longest noun stem and no limit on the number of noun stems. Notice that this phenomenon is simply a case of infinite agreement, that is, a case where the repeated elements are of boundless length and infinite number.

The above result increases to a significant degree the minimal strength which is proveably necessary for any general theory of linguistic structure. That is, by considering only the weakest of all possible formal requirements on grammars, that they enumerate the correct set of terminal strings, it can be shown that phrase structure systems which make no use of contexts are too weak. This is thus an extension of the kind of result achieved by Chomsky[140] when he showed that the Markov process or finite state conception of grammatical theory proposed by Hockett[141] and others was inadequate. The finite state theory is equivalent to allowing only rewrite rules of the forms A → B c, A → c. The result based on Mohawk shows that it is not enough to extend grammatical theory by allowing rules of the form A → B C.[142]

It is extremely important to note that there is no known way of enumerating the construction which shows Mohawk beyond CF-PSG description even by using CS-PSG rules. In fact it is, I think, almost certainly true that CS-PSG rules cannot enumerate natural languages. This has, however, not yet been proven. If it can, we will know that PSG cannot even meet the weakest possible grammatical requirements. Notice, incidentally, that the lack of a proof of the inadequacy of CS-PSG with respect to weak generative power in no way supports the validity of such devices. The burden of proof remains completely with the supporters of such systems. At the moment there are sets of sentences in natural languages which simply cannot be described in such terms in any known way regardless of how deviant or incorrect the SD are permitted to be. It should also be noted that PSG rules supplemented by discontinuous rewrite rules of any of the types discussed earlier in connection with discontinuities have in this regard the same status as CS-PSG. For these also there is no known derivation of constructions with infinite agreement.

It should perhaps be noted that it is possible to generate languages with the structure /γγ/ by using CS rules which make essential use of the power to permute and thus violate Condition (3).[143] Even by using this device, which would reduce the P-markers of Mohawk to absurdity, at least $6n^2$ rules are required to enumerate the construction, where n is the vocabulary involved in the equivalence. Furthermore, a minimum of six symbols must be mentioned in each rule. Since the relevant vocabulary in Mohawk is probably about a thousand, it follows that even this unacceptable rewrite grammar

requires on the order of six million rules mentioning thirty-six million grammatical symbols. This construction can be characterized by two simple transformations which provide correct SD.[144]

It should be emphasized that the construction which shows Mohawk beyond CF-PSG description and beyond any known type of CS-PSG description is not an isolated curiosity. It has been noted by Bar-Hillel and Shamir[145] that English constructions with respectively have the property /$\gamma\gamma$/. It might be argued, however, that these are rather peripheral. However, constructions of the form /$\gamma\gamma$/ are simply special cases of sequences in which there are separated strings of boundless length with infinite, overlapping dependencies. But it has been pointed out by Chomsky (personal communication) that English possesses instances of this, especially of the type /$\gamma\beta$/, γ and β necessarily distinct, in particular constructions containing comparatives and constructions containing so.[146]

It is, of course, logically possible that at least some other natural languages may be weakly generated by CF-PSG. This is, however, extremely unlikely. At any rate, just the fact that Mohawk cannot be is sufficient to show that general linguistic theory must contain more powerful mechanisms even apart from consideration of strong generative power. And this is unquestionable when SD and simplicity are considered.

CHAPTER 8: CONCLUSION

I. Substantive Conclusions

In the present study we have considered eight major formulations of American descriptive grammar. It should be clear that the views chosen for study embody ideas which are characteristic of modern American descriptive grammar in general. The overwhelming majority of actual linguistic descriptions written in modern times fall into the range of ideas found in the eight works considered in detail above. Many of these are actual applications of one or another of the systems we have discussed. I think that the views we have considered are fully representative of a whole genre of linguistic description. It is therefore not unfair to assert that any conclusions which hold for these eight formulations as a whole must hold for modern American conceptions of morphosyntactic description generally.

The eight views of grammar which were discussed fall naturally into two groups. The first, consisting of those formulated by Bloch, Wells, Harris, and Hockett were proposed before the development of TG. It was shown that, with the exception of discontinuities and the possible exception of unbounded branching in Wells' case, the conceptions of grammatical structure in these formulations are versions of PSG. The conception of SD in these views is a single P-marker, that is, a hierarchically structured labelled bracketing. Thus the claim that Chomsky's notion of PSG reasonably and correctly formalizes the immediate constituent approach to linguistic analysis as developed in America appears well-founded.[147] The only aspects of description utilized by these authors which are <u>not</u> formalized by PSG, namely, discontinuities and perhaps unbounded branching, do receive an apparently correct formalization in <u>TG</u> and this is indeed the only such formalization known.

The PSG conception of grammar, which is a natural formalization of the syntactic approach of American structural linguistics, is a quite reasonable theory of natural language which unquestionably formalizes many actual properties of human languages. However, if taken as a full theory of grammar, it is false, suffering from at

least those deficiencies noted in Chapter 7. The conception of both SD and of the rules which enumerate SD provided by general linguistic theory must be enriched if human languages are to be correctly characterized. TG is one such proposed enrichening which claims that the notion of SD must be broadened to include a set of P-markers for each sentence as well as a structure of transformations, these being provided by a widening of the class of grammatical rules to include those which map P-markers into P-markers as well as those which map strings of symbols into other strings.

We also considered four conceptions of grammar which are essentially post TG in development (or, in the case of tagmemics, perhaps simultaneous). It was noted, however, that there is nothing in these later views put forth by Pike, Lamb, Hockett, and Harris which avoids the problems of PSG. Like the pre TG views, these later formulations are also unformalized versions of PSG. Indeed, it is striking to note that for the most part they are in fact weaker versions than the earlier systems we considered because they appear, with the exception of Harris' string analysis, to be versions of CF-PSG, whereas it was clear that both Bloch's and Harris' early formulations were CS and this was not ruled out in Wells' work or in Hockett's earlier ideas. Thus, far from enrichening the theory of language, recent attempts to describe the form of grammar without transformations have actually impoverished linguistic theory by eliminating even that unformalized use previously made of contexts, and by imposing such principles as Condition (5) and Hockett's assumption that no atomic element may belong to more than one class. None of this later group of four grammatical formulations appears to be either as empirically adequate or as clearly stated as the ideas in Bloch's early approach to syntactic description. The claim on the part of several of the authors of these later views that their conceptions are reasonable alternatives to TG cannot be accepted. Despite its tentative character, there is no known alternative to TG. Certainly none of the views considered in this monograph are candidates for this status.

II. Methodological Conclusions

The grammatical ideas discussed in this paper are essentially versions of the theory of PSG. But it is evident that they are unformalized and unclear versions. A failure to make precise the notion of grammatical rule, which we have been forced to note with respect

to each view considered, is not a mere flaw in elegance. It is essentially connected to the fact that although American taxonomic linguists have been elaborating and applying the notions of PSG for several decades, the inadequacies of these have for the most part escaped detection. Only precise ideas are subject to real disconfirmation as well as to real confirmation. Only a precise conception of the form of grammatical rules makes it possible for a grammar to make a clear and thus testable claim about the sentences of individual languages. It is no accident, I think, that the developers of TG were the first to make precise the notions involved in PSG.

NOTES

1. By restricting attention to American writings in this study I do not wish to give the impression that I feel current European grammatical ideas lie beyond the generalizations reached here. Nor do I wish to give the impression that European conceptions are not worthy of study. Quite the contrary is the case on both counts. But considerations of space force some limitation to be made and I think it not unfair to say that American ideas have been in the forefront of the development of modern descriptive grammar. Furthermore, there is clearly enough similarity between typical American and European views to suggest that conclusions which hold for all of the former hold also for the latter. Thus although we cannot definitely assert that any conclusions which are reached here for American views must hold also for modern European conceptions of the form of grammar, it is reasonable to conclude that at least part of the burden of proof will lie with supporters of those European conceptions of grammar which are similar to the American views discussed in this monograph.

2. Zellig S. Harris, 'From Morpheme to Utterance', in Martin Joos, ed., Readings in Linguistics, (Washington, 1957); henceforth the latter work will be referred to as RIL; Methods in Structural Linguistics, esp. Chap. 16 (Chicago, 1951); Bernard Bloch, 'Studies in Colloquial Japanese II: Syntax', in RIL; Rulon Wells, 'Immediate Constituents', in RIL; Charles F. Hockett, 'Two Models of Grammatical Description', in RIL; Kenneth L. Pike, Language in Relation to a Unified Theory of the Structure of Human Behavior, Parts I, II, and III (Glendale, 1954, 1955, 1960); Robert E. Longacre 'String Constituent Analysis', Lg. 36.63-88 (1960); Sydney M. Lamb, Outline of Stratificational Grammar (Berkeley, 1962); Charles F. Hockett, 'Grammar for the Hearer', in Roman Jakobson, ed. Structure of Language and Its Mathematical Aspects, Proceedings of the 12th Symposium in Applied Mathematics (Providence, 1961); Harris, String Analysis of Sentence Structure (The Hague, 1962); Harris, 'Co-Occurrence and Transformation in Linguistic Structure', Lg. 33. 283-340 (1957); Noam Chomsky, The Logical Structure of Linguistic Theory (Cambridge, 1955); Syntactic Structures (The Hague, 1957); 'On the Notion 'Rule of Grammar'', in Structure of Language and Its Mathematical Aspects; 'A Transformational Approach to Syntax', in Proceedings of Third Texas Conference on Problems of Linguistic Analysis in English (Austin, 1962).

3. The terms 'transformation' and 'transformational rule' will be used in this paper in the sense described by Chomsky in the works noted in footnote 2

4. 'Three Models for the Description of Language', I. R. E. Transactions on Information Theory IT-2, No. 3 (1956); Syntactic Structures; 'On Certain Formal Properties of Grammars', Information and Control 2, No. 2 (1959); 'Formal Properties of Grammars',

to appear in Eugene Galanter, R. Duncan Luce, R. R. Bush, eds., Handbook of Mathematical Psychology, Vol. II. Henceforth the term 'phrase structure grammar' will be abbreviated PSG and the term 'transformational grammar' TG. These abbreviations and certain others introduced later will also be prefixed to the terms 'language', 'rule', 'description', etc., to indicate 'languages which can be generated by such grammars', 'rules contained in such grammars', 'descriptions with such grammars', etc.

5. 'The Logical Basis of Linguistic Theory' 510-11, to appear in Proceedings of the IXth International Congress of Linguists.

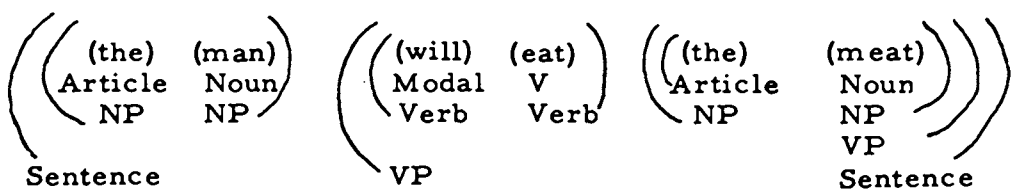
6. Robert B. Lees, review of Syntactic Structures. Lg. 33.385, (1957).

7. 'Logical Basis' 510.

8. A full theory of language would, furthermore, have to contain a device which enables the information provided by a grammar about completely well-formed strings to be extended to non well-formed utterances, to show the extent to which they are not well-formed, the types of deviation, relations between deviant utterances and nondeviant ones, etc. For some discussion see Chomsky, 'Some Methodological Remarks on Generative Grammar', Word 17. 219-39, (1961) and Jerrold Katz, 'Semi-Sentences', to appear in Jerrold Katz and Jerry Fodor, eds., Readings in the Philosophy of Language.

9. Notice that the question of simplicity is not merely a matter of elegance. Two or more grammars may be compatible with the same set of observed data and yet make logically incompatible claims about the language. Determination of simplicity is thus inherently bound up with the determination of truth.

10. For example, in terms of labelled parenthesizations the sentence in Diagram 1 could be represented as:



11. 'On the Notion 'Rule of Grammar''. Actually, the term 'PSG' is used in a systematically ambiguous way to refer both to the abstract underlying formal structures (discussed in greater detail below) and the diagrammatic representations of these in tree form.

12. This is asserted here without support. But if the general argument of the latter part of this study is correct, it will, in effect, have been demonstrated.

13. This argument does not depend on the infinitude of the class of sentences, only on the assumption that this class is too large to

be learned directly. Since, for example, the number of sentences in English of twenty words or less has been conservatively estimated at 10^{30} (there are about 3.15×10^{9} seconds per century) this assumption is beyond question. See George A. Miller, Eugene Galanter, and Karl H. Pribram, Plans and the Structure of Behavior 146-7 (New York, 1960).

14. Formally, the arrow represents a finite, two place relation.

15. These are essentially the so-called 'semi-Thue Systems' described in Martin Davis, Computability and Unsolvability, Chapter 6, (New York, 1958).

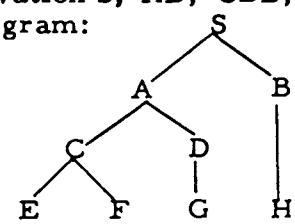
16. For a discussion see Davis, op. cit. 81-93, Chomsky, 'Formal Properties of Grammars'.

17. Notice that a language which is recursive is one for which there is a decision procedure or test for sentencehood. That is, given any arbitrary utterance, well-formed or not, there is an effective way of determining whether or not it is a sentence. If, however, a language consists of a set of strings which is recursively enumerable but not recursive, there is no decision procedure for sentencehood. That is, there will be an effective way to tell that any arbitrary presented sentence is a sentence, but no effective way to tell if an arbitrary presented nonsentence is or is not a sentence. Note that an 'effective' method is one which terminates is a finite number of steps.

18. 'Three Models'; Syntactic Structures; 'On Certain Formal Properties'; 'A Transformational Approach'.

19. This description must, of course, be made precise. In particular, the notion 'corresponding' must be carefully defined. For the precise statement, see Chomsky, Logical Structure, Chapter 6.

20. If P-markers are defined in terms of a single derivation, it will be necessary to define the relation 'is a' in a much more complicated way by making use of the notion 'traceable back' or the like. Note, for example, that in the derivation S, AB, CDB, CDH, CGH, EFGH, . . . which reduces to the diagram:



we must know that (EFG) is an A. Yet there is no pair of strings in the above derivation such that one can be formed from the other by substitution of (E F G) for A. Therefore, that (E F G) is an A must be formally characterized in a complex way by taking note of the fact that (E F G) is traceable through several different steps back

to A. There are many complexities in this course. In the full set of equivalent derivations, however, there will be two strings differing only in that A is replaced by (E F G).

An equally important reason which led Chomsky to take a P-marker to be the set of strings which occurs in the full class of equivalent derivations is that this makes possible an extremely natural and simple characterization of the notion 'linguistic level'. Within the general theory of linguistic structure developed by Chomsky, a linguistic level contains a Level-Marker (L-marker) for each sentence. Each L-marker is a set of strings in the 'vocabulary' of L which represents sentences. The SD of a sentence is the set containing all of its L-markers. In most cases an L-marker is simply a single representing string. Thus on the phonemic level the Phoneme-Marker consists of a single string of phonemes; on the level of phrase structure, L-markers consist of a larger set of strings; on the transformational level the T-markers consist of sets of strings, each representing one part of the transformational history (Concatenation of this level is interpreted as successive application. Thus the string ($T_{passive}$ $T_{question}$) is to be interpreted as the instruction: First apply $T_{passive}$ and then $T_{question}$ to the result.) The T-marker consists of a set with more than one member just in case there is a generalized transformation in the transformational history.

21. Furthermore, if Condition (2) c is not met there will be an unbounded number of lines possible in the derviation of any string and an infinite number of derviations for any terminal string. This would make the languages enumerated not recursive.

22. This is proven by Chomsky, 'On Certain Formal Properties of Grammars'.

23. 'On Certain Formal Properties of Grammars' 148.

24. We utilize here and below the notational convention of writing $A \rightarrow Z$ in X_Y for $XAY \rightarrow XZY$. We also utilize the notational devices of parentheses for optional elements and brackets for sets of mutually exclusive elements. Sometimes, instead of brackets, commas are used for disjunctions. These notational conventions are an important part of a metric of simplicity for grammars.

25. Condition (3) is a somewhat modified version of a restriction first suggested by R. J. Parikh, 'Language Generating Devices', Quarterly Progress Report No. 60, Research Laboratory of Electronics, Massachusetts Institute of Technology (1961). Parikh required that there be no rules of the form $XAY \rightarrow XBY$. This is too strong, however, and excludes many rules required for natural language description. Condition (3) and this restriction are probably identical in their effects on weak generative power.

26. 'Studies in Colloquial Japanese II: Syntax'.

27. RIL 160.

28. RIL.

29. To quote:

> 'We have given above a broad definition of the term construction; it admits as constructions a great many classes that are of no interest in the theory of ICs. But it is easy enough simply to ignore these uninteresting cases . . . ' RIL 194.

Thus Wells admits that his characterization does not distinguish between interesting and uninteresting cases and thus fails to characterize the notion being studied.

30. RIL 195.

31. Note that not every construction in Wells' sense is necessarily characterized by a single label in a P-marker. Some constructions will be represented by sequences of labelled elements, as the so-called 'actor action construction' which would presumably be represented by NP VP or the like. A similar remark holds for Hockett's use of 'construction' which is considered below.

32. RIL 198-9.

33. RIL 199.

34. A 'constitute' is defined by Wells, (RIL 188) as 'whatever sequence is constituted by two or more ICs . . . ' See the discussion below of Hockett's use of this term.

35. 'On Reformulating Transformational Grammars', Voprosy Jazykosznanija 10, No. 6 (1961).

36. In view of the fact that Wells' principle of multiple cuts is the closest, to my knowledge, that any modern linguist came before the formulation of TG to noting that sentences have infinite branching, it is odd that Robert E. Longacre, 'String Constituent Analysis', Lg. 36.63 (1960), should cite just this principle as typical of a reluctance to recognize more than two constituents. Longacre himself insists on branchings of higher degree than two. But in his formulation, as in that of tagmemic writers generally (see below), there is some finite upper bound on branching and the problem of infinite branching has not even been faced no less solved.

37. RIL 189.

38. 'From Morpheme to Utterance'; Methods in Structural Linguistics, especially Chapter 16.

39. Methods 263-4.

40. RIL 144-5.

41. G. H. Matthews has pointed out that R17 may make too strong a claim about Harris' equations. He notes that a formula like BC = A may assert only that there is some D such that: D → B C and D → A. This in no way affects the question of whether a PSG interpretation of Harris' formulas is possible.

42. RIL 147. Nothing is as yet known about the effects on weak generative power of restricting CS-PSG to contexts on one side only.

43. RIL 146.

44. Methods 266, footnote 267.

45. I am indebted to Noam Chomsky for this insight. A related point is that Harris needs the superscripts to indicate constraints which would be stated in a more ordinary PSG by means of intermediate constituents which he does not, and, because of his substitutibility requirements, probably cannot recognize. For example, he wishes to account for the fact that one can find boundless strings of adjectives but not of articles. He therefore must write AN = N but $TN^2 = N^3$. But in an ordinary PSG these constraints could be indicated by recognition of a higher order constituent NP and the rules NP → A NP; NP → T N.

46. Methods 267; RIL 146.

47. RIL 147.

48. A similar assumption to this one of Harris' was involved in Pike's original usage of the term 'tagmeme' which always had to have a single morpheme representative. However, this did not involve an empirical claim about language, as did Harris' system, because Pike allowed for higher order constituents besides tagmemes called 'hypertagmemes', 'syntagmemes', etc. which were not subject to this constraint. For some discussion see Velma Bernice Pickett, The Grammatical Hierarchy of Isthmus Zapotec 19, 90 (Baltimore, 1960).

49. RIL 389.

50. RIL 388.

51. Outline of Stratificational Grammar.

52. In fairness to Lamb it should be noted that this work is designed largely for classroom use.

53. I should not like it to be thought that this comment implies a derogation of the value of notational conventions for determining

simplicity. Indeed, these are an integral part of a theory of grammar and involve truth claims of the strongest sort. But conventions for defining the simplicity of rules of type X cannot influence the descriptive capacity or power of rules of type X.

54. Y. Bar-Hillel, M. Perles, and E. Shamir, On Formal Properties of Simple Phrase Structure Grammars 19-10 (Jerusalem, 1960).

55. Outline 6.

56. It is, of course, possible that some algorithm of SD assignment for rewrite rules exists which is not interfered with by deletions. But Lamb has not formulated any. The kinds which come to mind are complex and unnatural. Furthermore, these would require placing deletions in a separate component of the grammar. But this would mean that there are really two different structure assigning rules, one for deletions and one for ordinary PSG rules. To this extent, then, Lamb's system would resemble TG which has in the second component not only deletions but several other kinds of operations impossible in PSG. What would be missed in this approach to deletions is the great simplicity of TG and the fact that it incorporates the power of variables in rules while rewrite rules, with or without deletions, are limited to constants.

57. Outline 24.

58. Victor H. Yngve, 'A Model and an Hypothesis for Language Structure', Technical Report 369, Research Laboratory of Electronics, Massachusetts Institute of Technology (Cambridge, 1960).

59. Benjamin Elson and Velma Bernice Pickett, Beginning Morphology-Syntax (Santa Ana, 1960).

60. From the point of view of TG, the question of the degree of branching and the controversy over whether this may be more than two can be seen to be something of a pseudo issue. TG provide each sentence with a set of P-markers including, in particular, two which seem to be of real significance: the underlying P-marker provided by the PSG rules and the final derived P-marker which is the result of the application of all relevant transformations. In underlying P-markers the degree of branching is usually (though by no means always) two. In derived P-markers it seldom is chiefly because permutation transformations increase the degree of branching. For some discussion see Chomsky, 'On the Notion 'Rule of Grammar'' 23.

61. 'On Reformulating Transformational Grammars'.

62. Thus where an ordinary description will have rules like: VP → Verb NP, tagmemics will require the more complex: VP → P Object, P → Verb, Object → NP. Unfortunate redundancies due to

Condition (5) have in fact been noted by tagmemic writers, for example, by Viola Waterhouse, The Grammatical Structure of Oaxaca Chontal 119 (Baltimore, 1962).' But she mistakenly assumes this to be a mere notational inadequacy.

63. Language Part III 34.

64. This is, of course, not really the case and the whole assumption of a semantic interpretation of nonterminal elements is, I think, mistaken. Thus the subject in John pushes Bill has nothing very obvious in common semantically with the subject in John remembers his appointment, John received a blow, John suffered a defeat, etc. At this point, however, the tagmemic linguist is, of course, free to set up many different higher order tagmemes, including Subject versus Actor, etc. But this course, if followed to its natural conclusion, will lead to enormous complication of the description by requiring the recognition of dozens of grammatically equivalent constituents. First steps along this unhappy path have in fact been taken by Longacre, 'String Constituent Analysis' 70-73. Note that the semantic difference between subject in the above cases, as well as all analogous ones, is entirely due to the character of the verb. Thus marking it in terms of tagmemes is redundant even from the point of view of semantics. For a semantic conception which can take advantage of this fact and avoid redundancies see Jerrold Katz and Jerry Fodor, 'The Structure of a Semantic Theory', to appear in Language.

65. It might be argued that the proposed alternative is impossible in cases where there is free constituent order since in these subject and object might, for example, occur in identical positions with respect to the verb. This argument has not to my knowledge ever been given in favor of Condition (5). More importantly, as noted below, the argument for deriving free constituent order by transformation from one underlying constituent order is compelling. At any rate, it seems to me that the argument of footnote 64 is of overriding significance here.

66. 'Logical Basis' 528-9.

67. In our discussions of tagmemics we restrict attention to the notion 'tagmeme'. As indicated earlier, there are other related notions such as 'hypertagmeme', 'syntagememe', etc. However, as far as I can determine, the differences between these correspond only to the decision to refer to higher order constituents by a different name than lower order ones and to the highest order ones (sentence types) by a still different one (syntagmeme). This is of no theoretical significance in the present regard since the constituency of all such 'emes' is given by the same type of formulas.

68. This claim is supported by the fact that Elson and Pickett, in a work designed to familiarize students with tagmemic description, provide only formulas which have a CF interpretation.

69. Lawrence Clark, 'Sayula Popoluca Morpho-Syntax', IJAL 28.184-5 (1962).

70. 'Language Generating Devices'. Actually, Parikh's proof was given for CS grammars meeting the condition that there be no rule of the form XAY → XBY rather than Condition (3). But since the former is a stronger condition, the proof holds a fortiori for CS-PSG in the sense of the present work.

71. Pickett, The Grammatical Hierarchy of Isthmus Zapotec 20.

72. Pickett, op. cit. 40.

73. Language Part III 71.

74. Beginning Morphology-Syntax 63.

75. op. cit. 64.

76. This description is only suggestive and is not based on a serious study of Spanish grammar. In a real grammar it is likely that adjectives would be introduced transformationally. I do not think that this or any other difference would affect the present argument.

77. Note that in Spanish itself there must be at least one other Affix agreement rule, that between subject and predicate: Las alumnas son buenas 'the pupils (fem) are good'. Again a single transformation would suffice for a situation which would require a complex PSG description.

78. See Paul M. Postal, Some Syntactic Rules in Mohawk, esp. Chapter 3, Yale University Dissertation, (New Haven, 1962); 'Mohawk Prefix Generation', to appear in Proceedings of the IXth International Congress of Linguists.

79. Thus Pickett, Grammatical Hierarchy 42, does not even attempt to state concord by means of explicit rules. She instead merely lists the agreeing pairs.

It should be noted that this failure on the part of tagmemics to provide an explicit schema for grammatical rules has a natural result a failure to provide any method of determining the simplicity of descriptions. This causes a good deal of difficulty even in straightforward cases. Thus it was mentioned above that Elson and Pickett

hesitated to recognize gender morphemes in the Spanish noun. Yet from the point of view of the transformational rule of agreement this is an absolute necessity, even in cases where this is not morphophonemically transparent, as mujer, etc. This is a good illustration of the fact that the chief 'criterion' for grammatical analyses is provided by the character of the generative rules designed to enumerate them.

80. Morphophonemics of Modern Hebrew, University of Pennsylvania Masters Thesis (Philadelphia, 1951).

81. Harris, 'Discontinuous Morphemes', Lg. 21. 121-7 (1945) suggested that at least certain types of agreement could be dealt with by considering the discontinuous restricted elements as representatives of a single morpheme. This suggestion leaves unanswered the question with which we are here concerned, namely, what types of rules describe the strings in question. This is as much a problem if the strings are to be described in the morphophonemics as in the syntax. All the arguments for transformational description would still hold. Furthermore, this suggestion by Harris seems to make the claim which is, I think, not acceptable that no rule of the syntax ever need apply to the results of agreement rules. A similar claim is implied by the much stronger assumption made by Hockett, 'Linguistic Elements and their Relations', Lg. 37.51 (1961), equally unacceptable in my view, that all obligatory rules are nonsyntactic. In the above cases in Spanish this would mean, for example, that no rule of Spanish syntax ever need apply to the Affix constituents of the Article or Adjective constituents. While I do not know whether this result is acceptable, in the description of Mohawk (Some Syntactic Rules, Chapter 3; 'Mohawk Prefix Generation') the rule called 'T-reflexive' must apply to the output of the obligatory nominal-verb agreement rules and T-reflexive is unquestionably a syntactic rule.

82. Language Part III 36.

83. in Structure of Language and Its Mathematical Aspects. As indicated by the title, Hockett conceives of grammar in this article from the point of view of a listener attempting to determine the structure of presented sentences. This, however, in no way affects our attempt to interpret Hockett's views in the light of generative rules, which of course are entirely neutral as between speaker and hearer (generation is not production). Note that a linguist who claims, as does Hockett, to construct a device X which takes as input arbitrary sentences and gives as output their SD claims more than one who just claims to have a device Y which enumerates all sentences with their SD. To construct a 'grammar for the hearer' one must take a generative grammar which enumerates sentences with their SD and add to that an effective strategy for determining the SD, given the sentence as input,

84. A Course in Modern Linguistics, especially 157-65 (New York, (New York, 1958).

85. 'Grammar for the Hearer' 228.

86. An interesting attempt to define the notion 'modifier' in terms of the full theory of TG is given by Lees, 'The Grammatical Basis of Some Semantic Notions', in Report of the Eleventh Annual Round Table Meeting on Linguistics and Language Studies (Washington D. C., 1962). This is a significant attempt to characterize a grammatical notion in exactly the sense in which Hockett's usage of construction type is not, even if it turns out to be false. Notice that the question transformation proposed by Chomsky, Syntactic Structures 63-7, automatically reconstructs the relation between the two constructions of John is here and is John here while avoiding the consequence that a similar relation holds between any arbitrary pair of constructions.

87. 'Grammar for the Hearer' 228-30.

88. It is notable that in his A Course in Modern Linguistics 62, the consequences of the insistence on complete substitutibility are avoided by being vague. There he speaks of form classes as only having similar privileges of occurrence. Within the classificatory frame of reference he adopts, an attempt at precision leads to the unacceptable definiteness of the later version.

89. 'Grammar for the Hearer' 229.

90. Such putative members as dirty and clean fail because of the lack of *she will yellow up the room.

91. Which follows, of course, directly from Hockett's statement (3).

92. 'Grammar for the Hearer' 222.

93. Wells, RIL 189; Harris, Methods 265; Pike, Language Part I 151; C. E. Bazell, Linguistic Form 66 (Istanbul, 1953); Randolph Quirk, 'Substitutions and Syntactical Research', Archivum Linguisticum 10.40 (1958), to name only some.

94. 'Grammar for the Hearer' 229-30.

95. op. cit. 230.

96. Notice that TG provides a stronger account of ambiguity. In a TG it can be shown that two sentences are ambiguous not only because of either underlying or derived P-markers, but also because the same set of transformations was applied in a non equivalent order. See Chomsky, Syntactic Structures 88.

97. In proposed TG descriptions of English elements like and, or, etc. do have a special status because they are transformationally introduced constants and have no phrase structure origin. Thus some claim could be made that TG provides an account of a feature of language which Hockett feels is important but which CG fails to illuminate.

98. Interpretation of the notion 'reducibility' as strong equivalence entails a contradiction since Hockett admits that PSG does not permit discontinuities while CG may.

99. It might be claimed that the term 'PSG' is inappropriate here since Hockett is referring to Chomsky's earlier formulations in which permutations were not excluded, that is, in which there was no equivalent of Condition (3). However, although Condition (3) does affect weak generative power, this is of no importance in the present context since rewrite systems meeting Condition (2) are recursive regardless of whether Condition (3) (or its equivalent) is met or not.

100. It has been pointed out by Hilary Putnam, 'Some Issues in the Theory of Grammar', in Structure of Language and Its Mathematical Aspects (and by Lees, 'On Reformulating Transformational Grammars') that the power of TG to generate any recursively enumerable set must rightly be considered a flaw. Recently, G. H. Matthews, 'Analysis by Synthesis of Sentences of Natural Languages', to appear in Proceedings of the International Conference on Machine Translation of Languages and Applied Language Analysis, (Teddington, 1961), has proposed a condition which will render recursive the set of sentences enumerated by a TG by imposing restrictions on the way deletions may apply. There is too little known as yet to determine whether these restrictions are the correct ones. Notice that there is no reason to believe that even the restricted version of TG which enumerates only recursive sets is weakly equivalent to CS-PSG. Certainly Hockett's arguments do not in any way show this. As we note later, there are constructions in natural languages which have no known PSG derivation although they are easily described by TG with or without a restriction to recursive output. To show the equivalence of TG and CS-PSG, Hockett would at the least have to show that a CS-PSG derivation of these constructions is possible. Even this would not show the equivalence of any type of TG and CG, however, because as noted earlier the latter allows only CF-PSG rules.

101. Lees, 'On Reformulating Transformational Grammars', points out that both in the article we have been discussing and in 'Linguistic Elements and their Relations', Hockett misrepresents TG by taking optional singulary transformations like the passive in English to be typical of all transformations. Hockett shows how all such transformations can be eliminated by simply adding new 'constructions' to CG, that is, by recognizing new constituents in a PSG. But as Lees points out, the recursive power in a TG comes, not from optional singulary transformations like the passive, but rather from generalized or binary transformations like the nominalizations, complements

etc., which derive sentences from two or more underlying sentence structures. For these, Hockett's suggested replacement is simply impossible. Incidentally, Hockett does not consider what his suggested replacement of optional singulary transformations by 'constructions' would do to the simplicity of the grammar or the correctness of the SD, although in fact it was just such questions which motivated introduction of transformations like the passive in the first place.

102. String Analysis of Sentence Structure.

103. op. cit. 9.

104. op. cit. 12.

105. There is clearly an error here since the intention is that A is the constituent expansion.

106. op. cit. 13.

107. op. cit. 14.

108. Harris now seems to consider the word as the minimal element of constituent analysis. This change from the position of Methods is perhaps a reflection of a concentration on the study of written English. On page 12, in a footnote, Harris remarks that his approach in Methods and 'From Morpheme to Utterance' was 'An attempt to develop constituents systematically from single words . . .' However, as noted earlier, Harris' procedures in those works were based on morpheme not word classes.

109. There might conceivably be a procedure which distinguished these elements by a complex matching of sets of P-markers.

110. Hans G. Herzberger, 'The Joints of English' in Structure of Language and Its Mathematical Aspects 102-3.

111. 'The Grammatical Basis of Some Semantic Notions'.

112. For example, Lees notes that the elements (Be Ing) and (Have En), etc. are optional additions to the tense marker in the English auxiliary. Yet we should not wish to say that they were modifiers and the tense marker a head. There are many other such examples.

113. Notice that any successful characterization of these ideas will have to show that their explication is mechanically recoverable from the simplest grammar designed to enumerate the full set of sentences and set of correct associated SD. For example, if the notion of parenthesization is to be used for this purpose, it must be the case that for reasons quite independent of explicating 'modifier', the correct elements and no others must appear in parentheses in the simplest grammar.

114. It is not clear to me that Harris actually claims that string analysis is an independent alternative for syntactic description although he does claim (op. cit. 18) that it can describe all the sentences of a language (although he admits perhaps only with great complexity).

115. Chomsky points out (personal communication) that adjunction and substitution transformations may also yield discontinuities. Nonetheless, I think that the most typical cases are due to permutations. The unifying fact in all such cases is that there is an underlying P-marker in which the terminally nonadjacent elements are continuous constituents.

116. Roughly, this is to attach all branches broken under permutation to the next highest node, all other structure being preserved intact.

117. 'A Model and an Hypothesis for Language Structure'.

118. 'Discontinuity and Asymmetry in Phrase Structure Grammars', to appear in Information and Control.

119. As in the ad hoc diagrams given by Hockett, A Course in Modern Linguistics 155. Note that it is one thing to give some kind of diagramatic representation for discontinuities in a few sentences, quite another to give a general procedure which will assign such analyses to infinite sets of sentences, and still a third matter to show that these representations are in fact the correct ones.

120. Waterhouse, The Grammatical Structure of Oaxaca Chontal 45-6.

121. From the point of view of TG there is an interesting way of characterizing the tagmemic reaction to IC analysis of the ordinary, binary variety. It is possible that to a large extent the latter was attempting to characterize the underlying P-markers, while the former attempts nothing more than a characterization of final derived P-markers which have much less structure. Neither alone is adequate to handle discontinuities (among other phenomena) which can apparently only be correctly characterized in terms of the relations between a pair of P-markers.

122. These were discussed by Bar-Hillel in 'A Quasi-arithmetical Notation for Syntactic Description', Lg. 29.47-58 (1953). The proof is found in Y. Bar-Hillel, C. Gaifman, and E. Shamir, 'On Categorical and Phrase-Structure Grammars', The Bulletin of the Research Council of Israel 9f, No. 1 (1960).

123. Chomsky, 'Formal Properties of Grammars'; M. P. Schützenberger, 'Certain Elementary Families of Automata'.

Symposium on Mathematical Theory of Automata, Polytechnic Institute (Brooklyn, 1962).

124. A. G. Oettinger, 'Automatic Syntactic Analysis and the Pushdown Store', in Structure of Language and Its Mathematical Aspects; M. Sherry, 'The Identification of Nested Structures in Predictive Syntactic Analysis', to appear in Proceedings of the International Conference on Machine Translation of Languages and Applied Language Analysis; I. Rhodes, 'A New Approach to the Mechanical Translation of Russian', National Bureau of Standards Report No. 6295 (1959).

125. Maurice Gross, 'On the Equivalence of Models of Language Used in the Fields of Mechanical Translation and Information Retrieval', to appear in the Proceedings of the NATO Advanced Study Institute on Automatic Translation of Languages (Venice, 1962).

126. Lucien Tesnière, Elements de Syntaxe Structurale (Paris, 1959); D.G. Hays, 'Grouping and Dependency Theory', (The Rand Corporation, 1960).

127. 'Discontinuity and Assymetry in Phrase Structure Grammars'.

128. op. cit.

129. Pike, Language Part III 34; Hockett, 'Grammar for the Hearer' 230; Harris, String Analysis of Sentence Structure 12-13.

130. 'Grammar for the Hearer' 230, Language Part III 36.

131. Chomsky, 'On the Notion 'Rule of Grammar''; 'A Transformational Approach to Syntax'; 'The Logical Basis of Linguistic Theory'; Postal, Some Syntactic Rules'; 'Mohawk Prefix Generation'; Lees, 'Some Neglected Aspects of Parsing' to appear in Language Learning.

132. Chomsky, Syntactic Structures; Lees, review of Syntactic Structures; 'The Constituent Structure of Noun Phrases', American Speech, 36 (1961).

133. See the works cited in footnotes 131-2.

134. See the works cited in footnote 132.

135. Chomsky, 'Three Models'; 'Logical Basis'; Lees, review of Syntactic Structures; Postal, Some Syntactic Rules.

136. Chomsky, 'A Fragment of English Grammar', to appear; Emmon Bach, 'Subcategories in Transformational Grammars', to

appear in Proceedings of the IXth International Congress of Linguists; Paul Schachter, review of The Grammar of English Nominalizations, IJAL 28. 124-46 (1962).

137. Syntactic Structures 34.

138. Postal, 'Some Further Limitations of Phrase Structure Grammars', to appear in Jerrold Katz and Jerry Fodor, eds., Readings in the Philosophy of Language.

139. 'On Certain Formal Properties of Grammars'. Actually the proof was given for a special case but it immediately carries over to the more general formulation.

140. 'Three Models'.

141. A manual of Phonology, Memoir 11, IJAL (Baltimore, 1955).

142. PSG all of whose rules are of the form A → B C, or A → a are called 'normal' by Chomsky, 'Formal Properties of Grammars'. In his 'On Certain Formal Properties of Grammars' (where such grammars were called 'regular') it is proved that any language which has a CF-PSG at all has a normal grammar. Thus the restriction of P-markers to nodes with a maximum of two branches does not affect weak generative power. This result can probably be extended to CS-PSG although this is not yet proven.

143. This was shown by Chomsky in 'On Certain Formal Properties of Grammars'.

144. Postal, Some Syntactic Rules; 'Some Further Limitations of Phrase Structure Grammars'.

145. Y. Bar-Hillel and E. Shamir, 'Finite-State Languages: Formal Representations and Adequacy Problems', The Bulletin of the Research Council of Israel 8F, No. 3 (1960).

146. Thus one finds: I am more successful as a painter than John is as a writer, I am more successful as a painter than John is; but not: *I am more successful as a painter than John is as a painter. Similarly we find John drank the beer and so did Bill, but not *John drank the beer and so did Bill eat the meat.

147. This has sometimes been denied, for example, by Paul Garvin in his remarks at the 'Symposium on Transformational Theory' to appear in the Report of The Thirteenth Annual Round Table Meeting on Linguistics and Language Studies (Washington).

APPENDIX: HALLIDAY'S 'CATEGORIES OF THE THEORY OF GRAMMAR'

I. Introduction

In the earlier chapters of this work an attempt was made to demonstrate two basic points. First, that the syntactic ideas of representative examples of modern American descriptive linguistics may correctly be represented in terms of the notions of the theory of PSG, and second, that this theory must be regarded as incorrect. Since the above was written I have become familiar with M.A.K. Halliday's 'Categories of the Theory of Grammar'.[148] Although Halliday's work lies outside of the arbitrarily delimited field of discussion of the present inquiry, it is quite relevant to its concerns.

In a footnote on its first page Halliday claims that the rejection by American linguists of the approach of 'Bloomfieldian' linguistics in favor of the transformational views of Chomsky[149] is due in effect not to the inadequacy of the 'Bloomfieldian' views but rather to the fact that they have lacked 'theoretical foundation'. Halliday then apparently claims to provide this 'theoretical foundation' which presumably can preserve much of the pretransformational conception of language developed in modern linguistics, that is, 'taxonomic linguistics' in Chomsky's sense. This is not to say that everything Halliday writes is acceptable to American linguists or is believed by him to be acceptable to them. Quite the contrary. He even has a section where he lists what he considers to be the chief flaws of the 'Bloomfieldian' approach. But he obviously does feel that his ideas, which quite clearly are very similar to those held by many British linguists in the 'Firthian' tradition, provide an alternative to TG, an alternative which although distinct in many ways from 'Bloomfieldian' approaches is closer to the latter than to the approach suggested by Chomsky.

One disclaimer is necessary. The above remarks do not give an entirely accurate picture of Halliday's conception of his aims. He

apparently does not so much claim to give a particular theory alternative to TG but rather, in some sense, a general theory of language so broad that it in a way 'includes' TG, which in Halliday's view is concerned not with 'theory' but with 'description'. This position is quite incomprehensible to me since anyone who has followed the development of TG could not fail to have noticed that this is a quite general, language-independent theory[150] which intends to characterize the form of grammatical rules believed to occur in the grammars of all languages. This kind of theory is concerned with linguistic universals and is quite distinct from particular TG descriptions of individual languages which, of course, draw on the apparatus characterized in the general theory of TG. It thus seems proper to ignore this notion of Halliday's and treat his ideas as proposing a conception of the form of grammar and thus having a general logical status quite like that of the theory of TG. This appears to be perfectly compatible with the intentions of most of his writing.

I shall argue in this appendix that Halliday is partly correct. Without question his views are much closer to the kind of grammatical thought discussed in earlier chapters than they are to the transformational theory which has developed by pursuing the goals of generative grammar, goals quite foreign to Halliday's methodological framework. Indeed, Halliday's substantive views of grammar are so similar to those discussed above that they fit directly into the framework established in Chapters 1 and 2. Establishing this and drawing the appropriate conclusions from it are the goals of the sections that follow.

II. Basic Notions

Halliday makes at the outset an important distinction between linguistic theory, individual linguistic descriptions, and the presentation of either of these. He and we also are concerned with linguistic theory which, it is agreed, must provide the conceptual apparatus upon which individual descriptions of particular languages may draw.

For Halliday linguistic theory has several subdivisions, one of which is grammar, which concerns us here. Grammatical theory is for him characterized by four fundamental 'categories', namely, unit, structure, class, and system. He makes very strong empirical claims for these notions, namely: (all quotes from Halliday will be numbered for ease of reference) (1) "As the primary categories of

the theory, they make possible a coherent (emphasis mine: PMP) account of what grammar is and of its place in language, and a comprehensive description of the grammars of languages, neither of which is possible without them. "[151]

The category unit is set up (2) "to account for the stretches (of speech: PMP) that carry grammatical patterns. "[152] The units form a hierarchy, that is, a system of terms related along a single dimension by some form of logical precedence, which Halliday discusses in terms of 'highest' or 'lowest'. The scale along which the units are ranged is called rank. We are told: (3) "The relation among the units, then, is that, going from top (largest) to bottom (smallest), each 'consists of' one, or of more than one, of the unit next below (next smaller). "[153] 'Consists of' is thus a technical term.

The number of different units is, for Halliday, a fact particular to each language. It is universally required by him only that for each language there be at least two, sentence and morpheme. English grammar, he claims, has five: sentence, clause, group, word, and morpheme, in order of rank (technical term). Thus: (4) "So in the description of English the sentence consists of one or more complete clauses, the clause of one or more complete groups, the group of one or more complete words and the word of one or more complete morphemes. "[154]

The second category is structure. This is set up (5) "to account for likeness between events in successivity. "[155] A structure is made up of elements which occur in ordered relation. (6) "A structure is thus an arrangement of elements ordered in 'places'. "[156] Places are distinguished by order alone. And a structure is always a structure of a specified unit. Each unit in a description may display a range of possible different structures. The one requirement imposed by Halliday is that each unit (7) "must carry at least one structure that consists of more than one place. "[157] He claims that: (8) "Each place and each element in the structure of a given unit is defined with reference to the unit next below. Each place is the place of operation of one member of the unit next below, considered as one occurrence. Each element represents the potentiality of operation of a member of one grouping of members of the unit next below, considered as one item grouping. "[158] The item groupings mentioned are in fact discussed by Halliday in terms of the third category, class, to which we turn presently. From the remarks given Halliday draws the conclusion that the lowest rank unit, morpheme, has no structure.

The mode of actually specifying structures is given by Halliday as follows: (9) "In description, structures are stated as linear

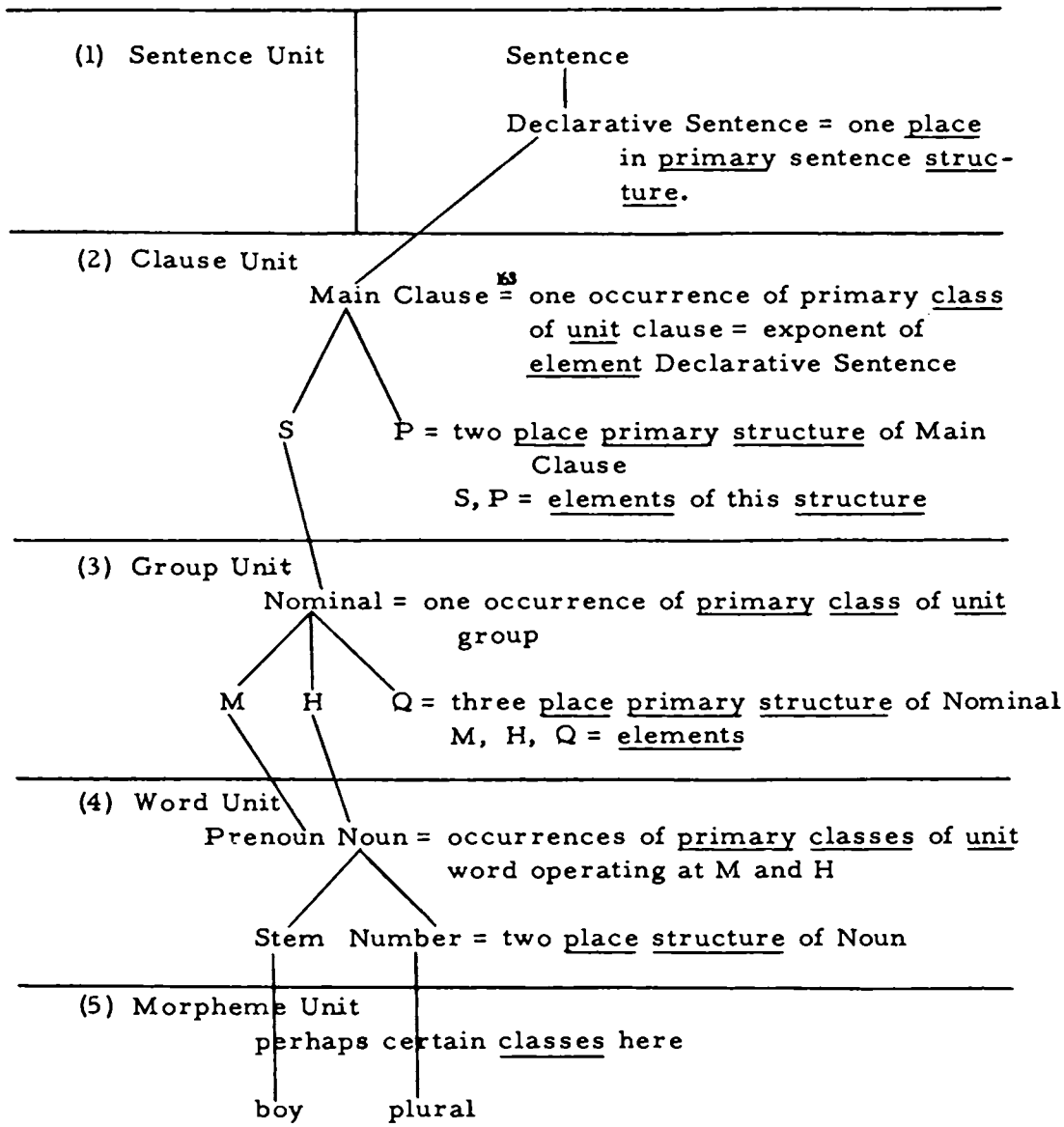
arrangements of symbols, each symbol (occurrence) standing for one place and each different symbol (item) standing for one element."[159] These ideas are illustrated only with the barest of examples. We are informed that in the description of English clause structure four elements are needed which have the widely accepted terms 'subject', 'predicator', 'complement', and 'adjunct' (abbreviated as are all other elements by their first letters, capitalized). All clause structures can, for Halliday, be stated as combinations of these four elements in different places. On the other hand, at the rank of the unit group the primary elements of structure of the nominal group are 'modifier', 'head', and 'qualifier'. Halliday notes that the possible primary nominal structures can be stated as a single formula where parentheses indicate optional occurrence, namely: (M)H(Q).

It was noted above that the places in structure are points of operation for groupings of members of the unit next below. (10) "To one place in structure corresponds one occurrence of the unit next below, and at each element operates one grouping of members of the unit next below."[160] The third basic category of the system in then introduced: (11) "The class is that grouping of members of a given unit which is defined by operation in the structure of the unit next above."[161] Finally, class, structure, and unit are related with the remark: (12) "Class, like structure, is linked to unit: a class is always a class of (members of) a given unit: and the class-structure relation is constant - a class is always defined with reference to the structure of the unit next above, and structure with reference to classes of the unit next below. A class is not a grouping of members of a given unit which are alike in their own structure. In other words, by reference to the rank scale, classes are derived 'from above' (or 'downwards') and not 'from below' (or 'upwards')."[162]

Before considering the final category of Halliday's approach to grammar, I shall try to give an interpretation in the generative terms formulated in Chapter 1 of the ideas considered thus far. Although the definitions of the three categories are quite vague and unclear (quotes (2), (5), and (11)) it is evident that unit, structure, and class all refer to elements of the SD of sentences and in fact to elements capable of representation with P-markers. The category unit is most closely related to the concept denoted by normal American descriptivist usage of 'level'. Thus note that where Halliday speaks of the different ranks of unit: sentence, clause, group, word, morpheme, Pike and his associates, for example, speak of sentence level tagmemes, clause level tagmemes, phrase level tagmemes, word level tagmemes, etc.

Interrelations between class, structure and unit may best be discussed in terms of the following illustrative, partly incomplete P-marker, described in Halliday's terms.

Diagram 26



It can be seen that the notion unit simply amounts to a division of constituents into sets, ranged from highest to lowest. Within each

unit there is in each position a single symbol, one class occurrence, which is dominated by a single element of structure at the highest unit and which dominates a string of symbols which represent its own primary structure. Each of these symbols in turn dominates one class occurrence of the next lowest unit, etc. This permits us to understand quote (12) which asserts that classes are defined with reference to the structure of the unit next above, and structure with reference to classes of the unit next below.

In terms of P-markers, the primary feature of the relation of structure and class is that each class occurrence is dominated by a single element of the structure of some higher unit class occurrence and that each element of every structure dominates one class occurrence of the next lowest unit (cf. quote (8)). Unfortunately, although this is the interpretation apparently intended throughout much of Halliday's discussion, it is not maintained consistently. This follows from the fact that Halliday allows subcategorization or subclassification within units. Thus not only does he discuss primary classes and primary structures, but also these may be broken down at the same unit into secondary structures and classes, and so on. Thus he claims[164] that although (i) all the ten houses on the riverside, and (ii) the finest old houses on the riverside, two exponents of the (class) nominal (of the unit) group have the same primary structure MHQ, they have distinct secondary structures still at group rank, namely, the first M is D_aD_bO, while the second M is D_bOE. Thus we would presumably find for (i) and (ii) the respective P-markers:

Diagram 27

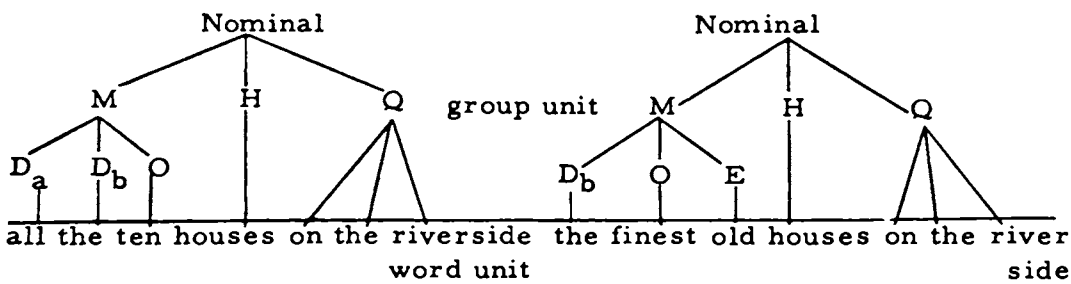


Diagram 28

But now it is evident that analyses such as this contradict quotes (8) and (10), since M is one place in the structure of the class occurrence, Nominal, of the unit group, but it does not correspond to one

occurrence of the unit next below (word), but rather to three such occurrences, and at M it is not true that there operates one grouping of members of the unit next below but rather three such groupings. At the very best then (8) and (10) are true only for the lowest or 'most delicate' differentiation of structure within a given unit. But most of the characterization of structure and class given by Halliday is in fact inconsistent with the notion of subclassification into primary, secondary, etc. structures within a unit and only makes sense if each class occurrence of unit X has only one level of structure in X. But this is, of course, quite incompatible with even the grossest conditions of empirical adequacy, given a restriction to five units (for English).

It is easy to appreciate the difficulty of maintaining logical consistency within an unformalized descriptive system like Halliday's. The fact that the descriptive categories discussed are almost never faced with any real data adds to the problem of maintaining conceptual clarity. But there is good reason to suspect that the particular contradiction just exposed is due not so much to these factors as to the fact that Halliday's system contains a wholely arbitrary, unjustified, and as far as I can see empirically completely unmotivated distinction, namely, that between divisions within a unit which are referred to as primary, secondary, etc. and divisions between units. This distinction is exactly paralleled by another equally arbitrary distinction, namely, the between rank and delicacy. Distinctions among units are referred to as those of rank, but within a unit subclassification is spoken of in terms of delicacy. No empirical facts, no arguments are offered to justify these distinctions. The only thing which appears to distinguish them is the peculiar idea, advanced wholely without support, that the number of possible divisions within a unit is both infinite and arbitrary. Thus in the case represented by Diagram 27 there are no reasons given by Halliday why the step from Subject to Nominal is one cross-unit while that between Nominal and M H Q is within a unit. In fact, the question of justifying such a distinction is not even raised. Halliday apparently feels that it is enough to call them by different names in different cases, viz. class names versus element of structure names.

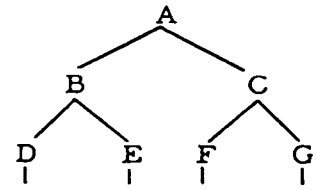
It might seem from the handful of English examples which are actually discussed by Halliday[165] that a possible ground for the distinction is this. All branchings, that is, cases where one symbol immediately dominates more than one other symbol are by definition intra-unit while configurations of the form:

Diagram 29

necessarily go from one element of a structure of unit X to a class of the next lower unit. However, this interpretation is not consistent with quotation (7) which clearly allows that some occurrences of the classes of some units will carry structures with only one place and thus will have unit internal representations like Diagram 29.

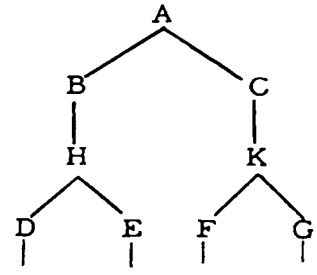
Notice that only the inconsistent recognition of secondary structures within a unit prevents Halliday's system from meeting Condition (5) which characterizes tagmemics. If we take quotes (8) and (10) literally, all P-markers in Halliday's terms must have in every position in each unit a single symbol, a class occurrence, dominating a string of other symbols, each of these in turn dominating a single symbol, a class occurrence, in the next lower unit. But this means that there are no P-marker configurations of the form:

Diagram 30



but only those of the form:

Diagram 31



Naturally, this restriction in Halliday's system will lead to the same kind of pointless redundancy as in tagmemic descriptions, that is, recognition of whole sets of otherwise unnecessary constituents, like S alongside Nominal, and H and K in Diagram 31. This similarity between Halliday's ideas and tagmemics is especially noteworthy since there is no indication that Halliday has either drawn on earlier tagmemic notions or is aware of any parallels between his work and the ideas of Pike and his associates

We can only conclude that the distinction between intra and inter-unit configurations has not been characterized by Halliday and is, in particular cases, completely arbitrary. In short, the distinction between element of structure and class, exactly parallel to the tagmemic distinction between slot and filler class, is unjustified. What is especially curious is that Halliday appears to lack any real motivation for this distinction and the associated acceptance of Condition (5) (of the type which, for example, motivated Pike's acceptance of this kind of notion). See, however, the discussion of grammatical relations in the following section. But if this distinction is removed the line between units dissolves.

It should be noted that one rational method which might be used to draw a distinction of levels or units within the branching structure of constituents, namely, a universal cross-language characterization of such concepts as clause, phrase, word, etc. is not attempted by Halliday and indeed is explicitly rejected as not part of his theory. Tagmemics, on the other hand, is quite interested in the universal specification of such concepts, even though, in my opinion, it has not moved very far in the direction of providing such.

We must conclude in other words that Halliday's work offers no grounds whatever for doubting the conclusions reached by Chomsky that: "On the level of constituent structure, then, each sentence of the language is represented by a set of strings, not by a single string as it is on the level of phonemes, morphemes, or words. Thus phrase structure, taken as a linguistic level, has the fundamentally different and nontrivial character which . . . is required for some linguistic level. We cannot set up a hierarchy among the various representations of 'the man hit the ball'; we cannot divide the system of phrase structure into a finite set of levels, ordered from higher to lower, with one representation for each sentence on each of these sublevels (emphasis mine: PMP). For example, there is no way of ordering the elements NP and VP relative to each other. Noun phrases are contained within verb phrases, and verb phrases within noun phrases in English.

Phrase structure must be considered as a single level, with a set of representations for each sentence of the language. "[166]

There is one final major category of Halliday's system which we have not discussed, namely system. This notion is concerned with subclassification, which has a direct, if often cumbersome description in PSG terms. Halliday remarks: (13) "If class 1 is the primary class (say of the group) operating at X in (clause) structure, and this has secondary classes 1.1, 1.2, and 1.3, then 1.1, 1.2, and 1.3 form a system of classes operating at X. "[167] The elements of systems are apparently referred to as terms. It seems, although here (264-7) the discussion becomes even more difficult to follow than elsewhere, that ultimately the notion of system is restricted to small grammatical classes with lexical sets discussed under the term lexis. This in no way effects PSG interpretation of Halliday's ideas and thus need not be discussed further. It is clear that in PSG terms system is simply a concept which refers to a set of elements dominated by certain other particular elements.

We can conclude this section of our discussion by noting that the ideas of unit, structure, class and system refer exclusively to aspects of SD capable of representation by P-markers. Further, these ideas are not developed consistently and involve unmotivated as well as unjustified distinctions. Certain of these ideas bear rather striking similarities to tagmemic formulations, although inconsistencies and lack of clarity on Halliday's part prevent them from being equivalents, even within the restricted range of P-marker characterization with which he is in effect concerned.

III. Some Further Inadequacies

It is evident that like other versions of taxonomic linguistics considered in earlier chapters, Halliday's ideas are restricted almost exclusively to the domain of the elements of SD of sentences and to the general (language independent) characterization of these. There is no attempt to specify the form of rules which must assign SD and their elements to infinite sets of sentences, consequently no attempt to specify the principles of structure assignment for these rules, their possible interrelations (ordering, mutual dependency, etc.); in short no real attempt to specify the notion 'grammar of a language in the sense of a theory which can describe exactly what utterances are and are not sentences of the language and which can assert what structure the sentences have. Given these limitations, it is not

surprising that Halliday says nothing about the crucial question of determining the simplicity of linguistic descriptions.[168] Thus although Halliday's notions of grammar quite clearly fall within the framework of PSG, he does not provide an account which is either clear or explicit enough to permit easy determination of the particular kinds of precise conditions his version of PSG would meet, were it specified exactly. We have seen, however, that insofar as his ideas are consistent (that is, insofar as only one level of structure analysis is permitted within each unit), his inter-unit - intra-unit distinction apparently places his ideas under Condition (5). There are a few other remarks which permit us to infer something about the rules which might be allowed. These have some further implications for the characterization of his account of SD.

Halliday claims the following: (14) "First, the theory allows for downward 'rank shift': the transfer of a (formal realization of a) given unit to a lower rank. Second, it does not allow for upward rank shift. Third, only whole units can enter into higher units. Taken together these three mean that a unit can include, in what it consists of, a unit of rank higher than or equal to itself but not a unit of rank more than one degree lower than itself; and not, in any case, a part of any unit."[169] Despite the difficulty of interpreting (14), it seems that what is intended here is to allow P-markers in which elements of higher units are dominated by those of lower units, thus sentences by noun phrases, clauses by verb phrases, etc. This is then just the kind of thing which Chomsky pointed out in the sections quoted earlier. Note that Halliday says, however, nothing about the kinds of rules which are to derive such results. It is unfortunate that at just the point where precise, formal statements about linguistic rules and their structure assigning properties are required to make a serious claim about language we find a one word metaphor, 'transfer', which is never defined, nor, as far as I can tell, even mentioned again.

Furthermore, downward rank shift must be restricted if false and in fact inconsistent (for Halliday) results are not to ensue. Obviously not every kind of embedding of elements of higher units in lower can be found. Clearly, for example, there can be no rank shift such that morphemes dominate sentences (this would contradict Halliday's stricture that morphemes have no structure); also it is clear that there will never be sentences or clauses dominated by elements of the word unit (despite some rash remarks sometimes made about polysynthetic languages). Halliday, however, provides no explicit restrictions on rank shift which would ensure that these intolerable outputs do not result. Elsewhere we are told that rank shift may be

recursive. This marks Halliday's recognition that embeddings are boundless without providing any insight into how grammars can characterize this boundlessness.

After this discussion of rank shift there is a footnote in which Halliday attempts to relate his discussion of rank shift to the distinction made by Yngve between 'regressive' and 'progressive' structures. Although Halliday does not explain the fact, this is a distinction between P-markers in which there are branchings of the different forms found in Diagrams 32 and 33, respectively:

Diagram 32

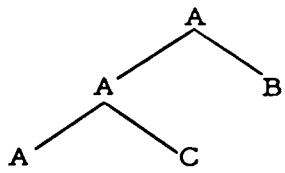
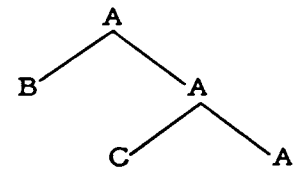


Diagram 33



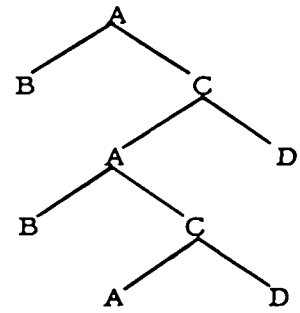
That is, the distinction between 'regressive' and 'progressive' is that between 'left branching' and 'right branching' structures.

Yngve claims[170] that in natural languages there is a limit to left but not to right branching.[171] That is, he claims there is a small finite bound on the number of symbols like A which may appear in P-markers like those of Diagram 32. Halliday says: (15) " Whether or not he (Yngve: PMP) is right in postulating a depth limit (of about 7) for 'regressive' structures, while allowing 'progressive' structures to be infinitely expanded, they do represent very different types of 'infiniteness' (emphasis mine: PMP) and are separately accounted for in the present theory, the former with, the latter without, rank shift. "[172] As with so many of the remarks made in this paper, no justification is offered for this claim, that is, for the assertion that in natural languages left branching constructions do not involve P-markers in which 'lower unit' elements dominate 'higher unit' ones while right branching ones do. Since he presents no evidence in favor of his claim, however, there is no reason to justify a contradictory one here, although I believe such is easily done.

Equally as important, the claim that left and right branching are 'very different types of infiniteness' is completely unsupported. Halliday offers no fact in favor of this view and indeed I know of none. Yngve's view that these two are very different follows not, as has

often been assumed, from any facts about infiniteness of constructions and finiteness of memory, but rather solely from the quite dubious assumption that in producing a sentence a speaker derives it 'from top to bottom', that is, by choosing higher level elements (S, NP, VP) before lower level ones (Noun, particular morphemes, etc.). What Halliday (and Yngve in the first place) have missed here is the very crucial and important set of differences between both infinite left and right branching and either self-embedding which yields P-markers like those of Diagram 34 or coordination.

Diagram 34



As has been pointed out several times before, coordination is incapable of correct description in PSG terms because of its boundless degree of branching. Neither left or right branching has this property. The existence of boundless self-embedding in a language, on the other hand, is exactly the property which renders a language beyond the possibility of description by finite state grammars. Neither left nor right branching has this property.[173]

Somewhat related to these curious ideas of Halliday's about a distinction between left and right-branching is his footnote about the way structures may be described. He notes: (16) " This formulaic presentation is useful as a generalized statement of an inventory of possible structures: a list H, MH, HQ, MHQ can be generalized as (M)H(Q). This particular instance is an oversimplification, since there may be more than one exponent of M and Q: the formula would then read (M..^{n})H(Q..^{n}), where ..^{n} allows infinite progression (not regression). "[174] It is interesting to speculate as to what precise kinds of rules and hence what exact claims about language are implied by notations of this sort for the description of recursion. First, there are two minor errors. Whatever else is intended, it is a mistake to have both M..^{n} and Q..^{n} since this implies the false claim

that both substructures always have the same number of places (thus no: the big book which John likes). One part of the notation must have $\ldots^{m}$. Further, given the use of the numerical indices, it is pointless to have the further conceptual apparatus of parentheses since the work of these is automatically given by having the indices take on O as values.

More importantly, what is the import of the claim that "$\ldots^{n}$ allows infinite progression (not regression)"? In Halliday's terms the claim is strictly meaningless since 'regressive' and 'progressive' are terms referring to P-markers and he has given no method for deriving P-markers. If, however, one utilizes the precise apparatus others have developed for assigning P-markers to infinite sets of strings, the notation being considered might be looked at as an abbreviation for the right sides of several PSG rules, namely those of roughly the form:

R76 NP → (M) X (Q)
R77 X → M X
R78 X → X Q
R79 X → H

Repeated application of such rules would produce right branching P-markers.

It seems, however, that this is not really Halliday's intention. Rather it is likely that he intended the strings of M and Q to be simply strung out, i.e. to be coordinated. But this is of course incompatible with the claim that the result is right branching. In fact, the coordinating result is right for one of the two structural interpretations which these kind of constructions have in English. However, this cannot be described in Halliday's terms and his ideas have the same unacknowledged or even considered inability to deal with coordination as do the versions of taxonomic linguistics considered in Chapter 4.

In a footnote[175] Halliday makes it clear that he recognizes discontinuous elements, in particular, discontinuous groups. He says nothing about the rules which are to assign such structure nor does he even make it clear how this notion relates to the rest of his own theoretical structure How, for example, are structures of discontinuous groups to be stated as sequences of symbols (quote (9))? Regardless of internal problems, it is clear that there is nothing in Halliday's approach to discontinuities which is in any way different from the unformalized approach to this matter in American descriptive linguistics. All the remarks of Chapter 5 hold without modification for Halliday's notions on this topic.

An important aspect of linguistic description concerns so-called 'grammatical relations' like those of 'subject', 'object', etc. As was pointed out earlier, tagmemics provides at best an obscure and redundant account of this feature of grammar by taking these relations to be extra constituents of a certain sort, the slots in which other constituents occur. Thus the tagmemic recognition of a subject constituent alongside Noun or Noun Phrase, etc. In fact, Halliday makes the same mistake, for many of his elements of structure are exactly superfluous constituents which attempt to represent grammatical relations. Thus he recognizes S (subject) alongside Nominal, exactly parallel to Pike's treatment. As was noted earlier, this is incorrect because it misses the 'relational' character of the notions involved and unnecessary because grammatical relations can be characterized in terms of subconfigurations of the ordinary constituents in P-markers without recognition of extra elements like S. However, the configurational technique for specifying grammatical relations is only an effective method in a TG in which there are abstract underlying P-markers. That is, there is a unique configuration of constituents associated with particular grammatical relations only in the underlying P-markers, not in the derived P-markers which represent the actual strings of words found in sentences. Thus in English the subject relation can be defined in terms of a subconfiguration NP VP in underlying P-markers. That is, the string of morphemes dominated by NP in this position bears this relation to the string dominated by VP. But if we attempt to use this characterization for derived P-markers there is no way to account for the presence of the subject relation in cases like is John eager to please, the killing of the tigers (ambiguous), John was seen by Mary, etc. But in a TG it can be shown that in the simplest grammar constructions like these are derived from underlying P-markers in which the appropriate elements are represented by configurations of the form NP VP. Thus the great effectiveness of the above configurational approach to grammatical relations is necessarily lost in exclusively PSG grammars since these can provide only P-markers equivalent to the final derived P-markers of TG descriptions.

The above is relevant because Halliday partially recognizes that grammatical relations have this abstract configurational character. Thus he says: (17) "There are some instances where an element of structure is identified as such solely by reference to formal sequence: where the element is defined by place stated as absolute or relative position in sequence. It is useful to indicate that here sequence is so to speak built in to structure, and this can be shown

by an arrow placed over the symbols for the elements concerned. For example, in English clause structure it is a crucial criterion of the element S that it precedes P in sequence: structures can be stated as $S\vec{P}CA$, $\vec{S}A\vec{P}A$, $A\vec{S}P$, etc. "[176] What Halliday misses here, of course, (besides the fact exemplified in the previous paragraph that the subject relation is not associated with any unique order of elements in actual sentences) is the fact that a configurational approach to grammatical relations renders his S symbol and all similar elements superfluous since everything that this represents can be given by a configurational account of the relations which Nominal bears to classes which operate at P (which is probably similarly redundant). Of course, this could only be effective for the full range of sentences if the relational characterizations are limited to underlying P-markers, with derived P-markers generated by transformations. We thus see that Halliday's arrow notation is interpretable in terms of the configurational characterization of grammatical relations and that if taken seriously it renders redundant his recognition of a wide class of elements like S, with the incorrect assimilation of grammatical relations to constituents which this implies.[177]

We have seen that everything which Halliday's unformalized and not always coherent approach to phrase structure permits us to infer about the rules lies completely within PSG interpretation with the one typical exception of discontinuities. There is one remark he makes which might be interpreted as counter to this conclusion. He says: (18) "... its (the technical concept 'consists of': PMP) realization in form varies between and within languages, and is stated of course in description. The possibilities are sequence, inclusion, and conflation. Thus if in a given instance a unit of one rank consists of two units of rank next below, these may appear as one following, interrupting, or overlaying the other. "[178] Only the first of these types of 'realization' is compatible with PSG characterization. However, although Halliday here allows the other two there is nothing in any of the other discussions or in the treatment of examples to suggest that he would make use of these possibilities in the syntax. Note, however, that inclusion or conflation do not affect PSG interpretation if they are restricted to the morphophonemics. Thus, for example, it is no argument against a PSG interpretation of a grammar of a Semitic language if that grammar recognizes discontinuous root and suffix morphemes. One need only assume that the syntax generates these as sequences of continuous, nonoverlapping elements and that the actual output is produced by morphophonemic rules, which need not, of course, be restricted to PSG form.[179] Since there is no evidence that Halliday intends to make use of any other

possibility than sequence in syntax, there is no reason to doubt the possibility of a PSG characterization of his notions of grammar on account of quote (18).

It should be emphasized that nothing in Halliday's discussions or treatment of examples suggests that a PSG interpretation of his ideas would require contexts in rules. Since there are no explicit statements on this topic, we follow reasonable practice and pick the weakest possible formal interpretation of the ideas under consideration. Our conclusion is then that, with the exception of an unformalized appeal to discontinuities, nothing in Halliday's approach to grammar would require apparatus beyond CF-PSG rules. Thus besides the general flaws of any PSG, and besides the many special flaws in Halliday's version we have discussed previously, any coherent interpretation of Halliday's system suffers as well from the defect of being provably incapable of enumerating all the sentences of human languages.

IV. Conclusion

Like most of the versions of American taxonomic linguistics considered in Chapter 4, Halliday's ideas about grammar amount only to an unformalized treatment of the SD of sentences. With the by now familiar exception of discontinuities, the account given is capable of interpretation in terms of a single P-marker per sentence. It was noted that Halliday's account of these ideas is quite inadequate, involving inconsistencies, a mass of unspecified points of great significance, several clearly incorrect claims, and generally such a lack of clarity of statement as to make its discussion rather unrewarding. Furthermore, all of this is correlated with the fact that the whole scheme is presented essentially without empirical support of the kind that could, for example, begin to justify the extremely strong claims of quote (1). It is important to point out that just as in the case of the more recent American views of constituent structure we discussed earlier, the material presented by Halliday appears to be much less adequate in terms of either internal coherence or empirical justification than that of the earliest extensive treatment of this topic, namely, Bloch's description of Japanese, in spite of the fact that the latter precedes the former by the better part of two decades.

A view of grammar which limits itself to informal discussion of SD and which ignores the question of a precise characterization of

linguistic rules can make only an extremely limited set of claims about language and leaves open an enormous range of questions which are necessarily not even raised. This raises the question in what sense such a view can be called a 'theory'. But further, the idea that the particular set of unformalized ideas about SD provided by Halliday, which are extraordinarily inadequate even within their own very limited domain of coverage, could be in any sense either an alternative to or inclusive of, the precise formal theory of TG is simply incomprehensible. It may be possible to propose a reasonable alternative to the theory of TG. An approach like that of 'Categories of the Theory of Grammar' does not even represent a step in this direction and indeed is a less adequate conception of grammar than the earliest attempts within modern linguistics to characterize the sentences of human languages in terms of the notions of constituent structure.

FURTHER NOTES

148. Word 17. 241-92 (1961).

149. I cannot imagine what gives Halliday the surely incorrect idea that a general abandonment of pre-transformational ideas in favor of the approach suggested by Chomsky is taking place in the United States. Neither the journals nor public meetings give this impression to me.

150. Cf., for example, the second sentence in Chomsky's Syntactic Structures.

151. ibid. 248

152. ibid. 251

153. ibid. 251

154. ibid. 253

155. ibid. 254

156. ibid. 255

157. ibid. 256

158. ibid. 256

159. ibid. 256

160. ibid. 259

161. ibid. 260

162. ibid. 261

163. Since Halliday does not provide even a tentatively complete description of one full sentence in this article, I have been forced to make up terms for the elements of the sentence unit, as well as those of the structure of the word unit. The description clearly follows his intention, however, insofar as this can be determined with the possible exception of taking one of the different sentence types to be an element of the structure of Sentence. Halliday himself says nothing about sentence types. I am not sure how to complete the morpheme unit.

164. ibid. 258

165. It is surprising that, in an article in which there is so little exemplification of theoretical ideas with real linguistic data, three full pages (277-80) are taken up with an analysis of eating in terms of his categories. Observe, incidentally, that any validity this analogy of the categorical analysis of eating and language might have would simply reveal the banality of the descriptive system by showing that it is too gross to distinguish speech from eating in any serious way.

166. Syntactic Structures 31-2.

167. ibid. 264

168. Although he admits (291) that linguistic theory must provide a practical evaluation procedure for grammars.

169. ibid. 251

170. A Model And An Hypothesis for Language Structure

171. For some discussion of the inadequacy of this view cf. Chomsky, "On the Notion 'Rule of Grammar'", and Chomsky and G. A. Miller, Finite Models of Language Users, to appear in the Handbook of Mathematical Psychology, Vol. II.

172. ibid. 251-2

173. For a discussion of all such matters relating to these branching types cf. Chomsky, "On the Notion 'Rule of Grammar'".

174. ibid. 257

175. ibid. 267

176. ibid. 257-8

177. It might be maintained that this remark is unfair to Halliday on the ground that he does differentiate elements like S from those like Nominal by calling the former 'elements of structure' and the latter 'classes'. But this is not a valid objection since, as we have seen, Halliday in fact provides no basis for this distinction and equally importantly whatever is implied in his account of elements like S does not provide the relational treatment which notions like subject require.

178. ibid. 251

179. This comment applies particularly to so-called 'phonological syntax', as developed for example by A. A. Hill, in his Introduction to Linguistic Structures, New York (1958). We have not considered this conception of grammar but would certainly claim a

PSG interpretation for it. At first glance this seems dubious because of the extensive recognition of 'supersegmental' morphemes of stress, pitch, juncture, etc. which do not occur in sequence with ordinary 'segmental' morphemes. However, it is easy enough to conceive of the syntax as generating all the morphemes sequentially with the multidimensional 'imposition' of the 'supersegmentals' 'on' ordinary morphemes resulting in the morphophonemics.

Notice, incidentally, that an argument like that referred to in the preceding paragraph cannot show that a TG may be reformulated as a PSG with more extensive morphophonemics (as suggested in effect by Hockett, 'Linguistic Elements as their Relations', Lg. 37.51-2 (1961)) for essentially the reasons noted in footnote 101.

BIBLIOGRAPHY

Bach, Emmon. "Subcategories in Transformational Grammars", to appear in Proceedings of the IXth International Congress of Linguists.

Bar-Hillel, Y. "A Quasi-arithmetical Notation for Syntactic Description", Language XXIX.

Bar-Hillel, Y., Gaifman, C., and Shamir, E. "On Categorical and Phrase-Structure Grammars", The Bulletin of the Research Council of Israel 9f, No. 3.

Bar-Hillel, Y., Perles, M., and Shamir, E. On Formal Properties of Simple Phrase Structure Grammars; Jerusalem, 1960.

Bar-Hillel, Y., and Shamir, E. "Finite-State Languages: Formal Representations and Adequacy Problems", The Bulletin of the Research Council of Israel 8F, No. 3 (1960).

Bazell, C.E. Linguistic Form; Istanbul, 1953.

Bloch, Bernard. "Studies in Colloquial Japanese II: Syntax", Readings in Linguistics; Washington, 1957.

Chomsky, N. and Miller, G. A. "Finite Models of Language Users", to appear in the Handbook of Mathematical Psychology, Vol. II, Luce, Galanter and Bush editors.

Chomsky, Noam. Morphophonemics of Modern Hebrew, University of Pennsylvania Masters Thesis; Philadelphia, 1951.

——— The Logical Structure of Linguistic Theory' Cambridge 1955.

——— "Three Models for the Description of Language", I.R.E. Transactions on Information Theory IT-2, No. 3 (1956).

——— Syntactic Structures; The Hague, 1957.

——— "On Certain Formal Properties of Grammars", Information and Control II, No. 2 (1959).

——— "Some Methodological Remarks on Generative Grammar", Word XVII (1959).

——— "On the Notion 'Rule of Grammar' ", Structure of Language and Its Mathematical Aspects, Proceedings of the 12th Symposium in Applied Mathematics; Providence, 1961.

——— "A Transformational Approach to Syntax", Proceedings of the Third Texas Conference on Problems of Linguistic Analysis in English; Austin, 1962.

——— "Formal Properties of Grammars", to appear in Eugene Galanter, R. Duncan Luce, R.R. Bush: Handbook of Mathematical Psychology, Vol. II.

——— "The Logical Basis of Linguistic Theory", to appear in Proceedings of the IXth International Congress of Linguists.

——— "A Fragment of English Grammar", to appear.

Clark, Lawrence. "Sayula Populuca Morpho-Syntax", IJAL 28. 184-5, (1962).

Davis, Martin. Computability and Unsolvability; New York, 1958.

Elson, Benjamin and Pickett, Velma Bernice. Beginning Morphology-Syntax; Santa Ana, 1960.

Garvin, Paul. "Symposium on Transformational Theory" remarks, to appear in the Report of the Thirteenth Annual Round Table Meeting on Linguistics and Language Studies; Washington.

Gross, Maurice. "On the Equivalence of Models of Language Used in the Fields of Mechanical Translation and Information Retrieval", to appear in the Proceedings of the NATO Advanced Study Institute on Automatic Translation of Languages; Venice, 1962

Halliday, M.A.K. "Categories of the Theory of Grammar", Word XVII. 241-92 (1961).

Harris, Zellig S. "Discontinuous Morphemes", Language XXI (1945)

——— Methods in Structural Linguistics; Chicago, 1951.

——— "Co-Occurrence and Transformation in Linguistic Structure", Language XXXIII (1957).

——— "From Morpheme to Utterance", Martin Joos: Readings in Linguistics; Washington, 1957.

——— String Analysis of Sentence Structure; The Hague, 1962.

Hays, D.G. "Grouping and Dependency Theory"; The Rand Corporation, 1960.

Herzberger, Hans G "The Joints of English", Structure of Language and Its Mathematical Aspects.

Hill, A.A. Introduction to Linguistic Structures, New York (1958).

Hockett, Charles F A Manual of Phonology, Memoir 11, IJAL; Baltimore, 1955.

——— "Two Models of Grammatical Description", Readings in Linguistics; Washington, 1957.

——— A Course in Modern Linguistics; New York, 1958.

——— "Linguistic Elements and their Relations", Language XXXVII (1961).

——— "Grammar for the Hearer", Roman Jakobson: Structure of Language and Its Mathematical Aspects, Proceedings of the 12th Symposium in Applied Mathematics; Providence, 1961.

Katz, Jerrold J. "Semi-Sentences", to appear in Jerrold J. Katz and Jerry A. Fodor: Readings in the Philosophy of Language.

Katz, Jerrold J. and Fodor, Jerry A. "The Structure of a Semantic Theory", to appear in Language.

Lamb, Sydney M. Outline of Stratificational Grammar; Berkeley, 1962.

Lees, Robert B. Review of Syntactic Structures, Language XXXIII (1957).

——— "On Reformulating Transformational Grammars", Voprosy Jazykoznanija X, No. 6 (1961).

——— "The Constituent Structure of Noun Phrases", American Speech, Vol. XXXVI (1961).

——— "The Grammatical Basis of Some Semantic Notions", Report of the Eleventh Annual Round Table Meeting on Linguistics and Language Studies; Washington, D.C., 1962.

——— 'Some Neglected Aspects of Parsing' to appear in Language Learning.

Longacre, Robert E. "String Constituent Analysis", Language XXXVI (1960).

Matthews, G.H. "Analysis by Synthesis of Sentences of Natural Language", to appear in Proceedings of the International Conference on Machine Translation of Languages and Applied Language Analysis; Teddington, 1961.

——— "Discontinuity and Asymmetry in Phrase Structure Grammars", to appear in Information and Control.

Miller, George A., Galanter, Eugene, and Pribram, Karl H. Plans and the Structure of Behavior; New York, 1960.

Oettinger, A.G "Automatic Syntactic Analysis and the Pushdown Store", Structure of Language and Its Mathematical Aspects.

Parikh, R.J. "Language Generating Devices", Quarterly Progress Report No.60, Research Laboratory of Electronics, Massachusetts Institute of Technology (1961).

Pickett, Velma Bernice. The Grammatical Hierarchy of Isthmus Zapotec; Baltimore, 1960.

Pike, Kenneth L. Language in Relation to a Unified Theory of the Structure of Human Behavior, Parts I, II, and III; Glendale, 1954, 1955, 1960.

Postal, Paul M. Some Syntactic Rules in Mohawk, Yale University Dissertation; New Haven, 1962.

——— "Some Further Limitations of Phrase Structure Grammars", to appear in Jerrold J. Katz and Jerry A. Fodor: Readings in the Philosophy of Language.

——— "Mohawk Prefix Generation", to appear in Proceedings of the IXth International Congress of Linguists.

Putnam, Hilary. "Some Issues in the Theory of Grammar", Structure of Language and Its Mathematical Aspects.

Quirk, Randolph. "Substitutions and Syntactical Research", Archivum Linguisticum X (1958).

Rhodes, I. "A New Approach to the Mechanical Translation of Russian", National Bureau of Standards Report No. 6295 (1959).

Schachter, Paul. Review of The Grammar of English Nominalizations, IJAL 28 (1962).

Schützenberger, M.P. "Certain Elementary Families of Automata", Symposium on Mathematical Theory of Automata, Polytechnic Institute; Brooklyn, 1962.

Sherry, M. "The Identification of Nested Structures in Predictive Syntactic Analysis", to appear in Proceedings of the International Conference on Machine Translation of Languages and Applied Language Analysis.

Tesnière, Lucien. Elements de Syntaxe Structural; Paris, 1959.

Waterhouse, Viola. The Grammatical Structure of Oaxaca Chontal; Baltimore, 1962.

Wells, Rulon. "Immediate Constituents", Readings in Linguistics; Washington, 1957.

Yngve, Victor H. "A Model and an Hypothesis for Language Structure", Technical Report 369, Research Laboratory of Electronics, Massachusetts Institute of Technology; Cambridge, 1960.

VOLUMES IN PRINT:

12. Robert B. Lees, The Grammar of English Nominalizations (xliv + 205 pp., 1966, reprint) $4.00

15. Robert Bruce Inverarity, Visual Files Coding Index (viii + 185 pp., 1960) $3.50

16. Isabella Y. Yen, A Grammatical Analysis of Syau Jing (xii + 159 pp., 1960) $4.00

24. Andreas Koutsoudas, Verb Morphology of Modern Greek (viii + 72 pp., 1962) $3.00

25. Joseph H. Greenberg, The Languages of Africa (viii + 177 pp., 1966, reprint) $5.00

26. Harold E. Driver and Wilhelmine Driver, Ethnography and Acculturation of the Chichimeca-Jonaz of Northeast Mexico (x + 265 pp., 1963) $4.00

27. Francis W. Gravit and Albert Valdman, Eds., Structural Drill and the Language Laboratory (xii + 224 pp., 1963) $4.00

28. V. S. Rastorgueva, A Short Sketch of Tajik Grammar (Trans. & Ed. by Herbert H. Paper) (xii + 110 pp., 1963) $3.00

29. V. S. Rastorgueva, A Short Sketch of the Grammar of Persian (Trans. by Steven P. Hill & Ed. by Herbert H. Paper) (x + 79 pp., 1964) $3.00

30. Paul Postal, Constituent Structure: A Study of Contemporary Models of Syntactic Description (viii + 122 pp., reprint, 1968) $6.00

31. Fred W. Householder, Kostas Kazazis, and Andreas Koutsoudas, Reference Grammar of Literary Dhimotiki (viii + 122 pp., 1964) $4.00

32. Norman Balfour Levin, The Assiniboine Language (xiii + 166 pp., 1964) $4.00

33. D. A. Shafeev, A Short Grammatical Outline of Pashto (Trans. and Ed. by Herbert H. Paper) (x + 90 pp., 1964) $3.00

34. Ilse Lehiste, Acoustical Characteristics of Selected English Consonants (xii + 197 pp., 1964) $5.00

35. V. I. Abaev, A Grammatical Sketch of Ossetic (Ed. by Herbert H. Paper and Trans. by Steven P. Hill) (x + 133 pp., 1964) $3.00

36. Joshua A. Fishman, Yiddish in America (viii + 96 pp., 1965) $3.00

37. Marvin I. Herzog, The Yiddish Language in Northern Poland (xxx + 324 pp., 1965) $5.00

38. G. Kingsley Noble, Proto-Arawakan and Its Descendants (x + 132 pp., 1965) $3.00

39. Charles T. Scott, Persian and Arabic Riddles: A Language-Centered Approach to Genre Definition (viii + 13 pp., 1965) $4.00

40. Edward W. Najam (Ed.), Language Learning: The Individual and the Process (xiv + 274 pp., 1966) $4.50

41. Edward H. Bendix, Componential Analysis of General Vocabulary: The Semantic Structure of a Set of Verbs in English, Hindi, and Japanese (x + 190 pp., 1966) $4.50

42. Elizabeth Bowman, The Minor and Fragmentary Sentences of a Corpus of Spoken English (xii + 68 pp., 1966) $3.00

43. Robbins Burling, Proto Lolo-Burmese (vi + 101 pp., 1966) $4.00

44. Stanley Lieberson, (Ed.), Explorations in Sociolinguistics (vi + 204 pp., 1967) $4.50

45. Madeleine Mathiot, An Approach to the Cognitive Study of Language (16 + 224 pp., 1968) $4.00

All correspondence and orders from the
United States of America and from Canada
should be addressed to
Publications, Research Center for the Language Sciences,
Patton House, Indiana University, Bloomington, Indiana 47401.

Orders from all other countries should be sent to
Mouton & Company, Publishers, The Hague, The Netherlands.